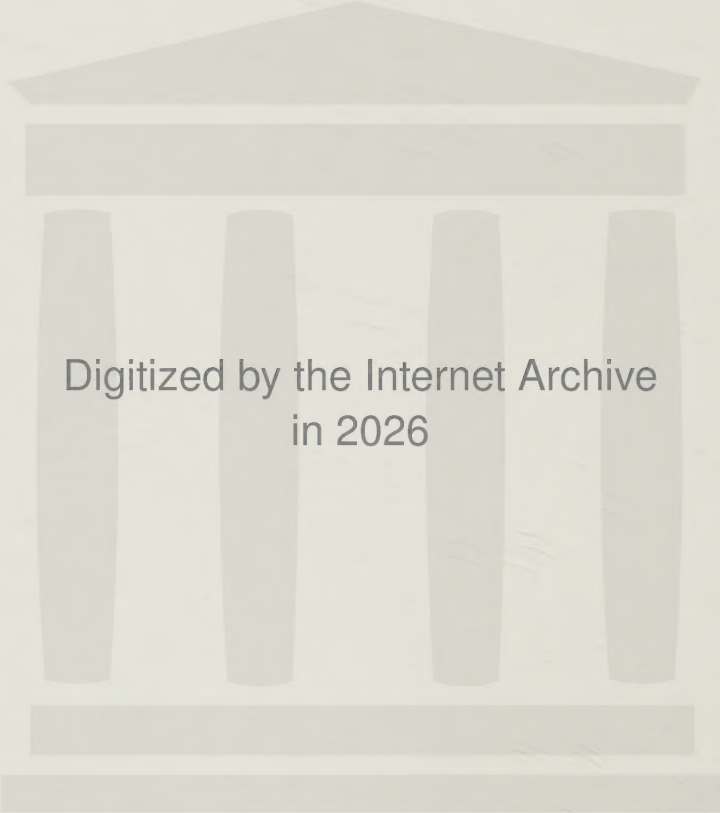


BEITRÄGE ZUR THEORIE
EMPIRISCHER PROZESSE,
INDIZIERT NACH
VAPNIK-CHERVONENKIS-KLASSEN

Wilhelm Schneemeier



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41547927_0023

BEITRÄGE ZUR THEORIE EMPIRISCHER PROZESSE,
INDIZIERT NACH VAPNIK-CHERVONENKIS-KLASSEN

DISSERTATION

zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften

der Fakultät für Mathematik

der Ludwig-Maximilians-Universität München

vorgelegt von

Wilhelm Schneemeier

eingereicht

am 14. Mai 1986

1. Berichterstatter: Prof. Dr. P. Gänßler, München
 2. Berichterstatter: Prof. Dr. H.G. Kellerer, München
- Tag der mündlichen Prüfung: 17.11.1986

INHALTSVERZEICHNIS

Einleitung	1
------------	---

Kapitel I: Eine Charakterisierung der L_b -Konvergenz von Zufallselementen in Funktionenräumen

1.1 Grundbegriffe und Definitionen	6
1.2 Ein fundamentales Approximationslemma	8
1.3 Charakterisierung der relativen L_b -Folgen- kompaktheit anhand des Stetigkeitsmoduls	13
1.4 Charakterisierung der L_b -Konvergenz und hinreichende Kriterien	23

Kapitel II: Ein funktionaler Grenzwertsatz für den "gebootstrappten" empirischen C-Prozeß, indiziert nach Vapnik-Chervonenkis-Klassen von Mengen

2.1 Stochastische Ungleichungen	33
2.2 Relative L_b -Folgenkompaktheit, L_b -Konvergenz sowie Glivenko-Cantelli-Aussagen für Dreiecks- schemata	43
2.3 Der gebootstrappte empirische C-Prozeß	49
2.4 Bemerkungen zum Fall gewichteter empirischer C-Prozesse	55

Kapitel III: L_b -Aussagen und Glivenko-Cantelli-Resul- tate für U-Prozesse, die indiziert sind nach Vap- nik-Chervonenkis-Klassen von Mengen

3.1 Definition des U-Prozesses und elementare Eigenschaften	57
--	----

3.2 Stochastische Ungleichungen	61
3.3 L_b -Konvergenz und Glivenko-Cantelli-Theorem	70
<u>Kapitel IV: Der gewichtete U-Prozeß auf der reellen</u>	
<u>Achse</u>	
4.1 Grundbegriffe und Definitionen	76
4.2 Stochastische Ungleichungen	77
4.3 L_b -Konvergenz des gewichteten U-Prozesses	81
4.4 Raten im Glivenko-Cantelli für den U-Prozeß auf der reellen Achse im i.i.d.-Fall	88
Literaturverzeichnis	96

EINLEITUNG

Ausgangspunkt dieser Arbeit waren die Untersuchungen zur L_b -Konvergenz in GÄNSSLER(1983). Dieses Modell ist in zweierlei Hinsicht von besonderem Interesse: Zum einen ist es geeignet, schwache Konvergenz von Zufallselementen zu formulieren, die sich in nicht notwendig separablen metrischen Räumen realisieren und daher im allgemeinen nicht meßbar sind bezüglich der zugehörigen Borel- σ -Algebra. Zum anderen vermeidet es den Begriff der δ -Straffheit, wie er in DUDLEY(1978) (vgl. hierzu auch GÄNSSLER(1983), S.58) Verwendung findet.

Wir gehen zunächst von folgender Situation aus: Sei (X, s) ein totalbeschränkter halbmetrischer Raum und $(X_n)_{n \in \mathbb{N}}$ eine Folge von Zufallselementen in S (wobei $S_0 \subset S \subset \ell^\infty(X)$; zur Definition vgl. 1.1.1). Unter geeigneten Meßbarkeitsvoraussetzungen (M1) und (M2) sind die $X_n, n \in \mathbb{N}$, meßbar bzgl. einer σ -Algebra A auf S mit $B_b(S) \subset A \subset B(S)$ (zur Definition vgl. 1.1), d.h. Theorem 11a) in GÄNSSLER(1983) sichert sofort eine Charakterisierung der relativen L_b -Folgenkompaktheit von $(X_n)_{n \in \mathbb{N}}$ (wobei wir per Definition verlangen, daß mögliche Grenzelemente sich auf dem separablen Raum S_0 konzentrieren).

Die Frage ist nun, ob sich ein analoges Resultat zu Theorem 1.17 in GÄNSSLER-HÄUSLER-SCHNEEMEIER(1984) (dort wird die klassische Situation $X=[0,1]$, $S_0=C([0,1])$, $S=D([0,1])$)

betrachtet) auch in unserem sehr viel allgemeineren Fall beweisen läßt. Es stellt sich heraus, daß dies sehr wohl möglich ist anhand geeigneter Modifikationen des dortigen Beweises in Verbindung mit dem Approximationslemma 1.2.1 der vorliegenden Arbeit. Die daraus resultierende Charakterisierung der relativen L_b -Folgenkompaktheit durch den Stetigkeitsmodul stellt das erste Hauptergebnis (Theorem 1.3.4) dar und verallgemeinert insbesondere Theorem B in GÄNSSLER(1983) bzw. Theorem 2.12 in GINÉ-ZINN(1984). Wie in GÄNSSLER-HÄUSLER-SCHNEEMEIER(1984) erhält man aus 1.3.4 sofort eine Charakterisierung der L_b -Konvergenz anhand der Konvergenz der endlichdimensionalen Randverteilungen und der Bedingungen (iii)(A) bzw. (iii)(B). Dabei besteht im allgemeinen die Schwierigkeit lediglich in der Sicherstellung von (iii)(A), wofür wir mit 1.4.3 hinreichende Bedingungen liefern, die in dieser allgemeinen Situation neu sind und bei deren Beweis das bekannte "Chaining Argument" sehr wesentlich eingeht.

Insbesondere lassen sich diese Ergebnisse anwenden, falls X eine Vapnik-Chervonenkis-Klasse ist und die Zuwächse von $X_n, n \in \mathbb{N}$, einer stochastischen Ungleichung vom Bernsteintyp genügen (vgl. Bemerkung 1.4.5).

Genau diese Situation treffen wir in Kapitel II an, wo wir für den "gebootstrappten" empirischen C -Prozeß L_b -Aussagen und Glivenko-Cantelli-Resultate beweisen, wobei C eine Vapnik-Chervonenkis-Klasse von Mengen ist (von der wir Abzähl-

barkeit verlangen, um der Einfachheit halber zusätzliche Meßbarkeitsvoraussetzungen zu vermeiden; vgl. hierzu die Bemerkung zu Beginn von Kapitel II). Durch bedingte Betrachtungsweise führt dies auf das Problem der Herleitung entsprechender Aussagen für empirische C-Prozesse, die auf beliebigen Dreiecksschemata von zeilenweise unabhängigen Zufallselementen basieren. Wegen 1.4.3 reduziert sich dies auf die Untersuchung "kleiner" Mengen, wofür Lemma 2.1.1 den entscheidenden Zugang liefert – ein Ergebnis, das in dieser Form neu ist und im Vergleich zu entsprechenden Aussagen in der Literatur sehr viel elementarer bewiesen wird (vgl. Bemerkung 2.1.3). Die auf diese Weise erzielten Bootstrap-Ergebnisse stellen im übrigen eine Verallgemeinerung dar von GÄNSSLER(1985), wo die i.i.d.-Situation abgehandelt wird. Den Abschluß des Kapitels II bilden schließlich einige neue Ergebnisse zu gewichteten "gebootstrappten" empirischen C-Prozessen.

Aufbauend auf entsprechenden stochastischen Ungleichungen für den empirischen C-Prozeß (2.1.4(ii) und 2.1.7) untersuchen wir in Kapitel III den sogenannten empirischen U-Prozeß, der wieder indiziert ist nach einer Vapnik-Chervonenkis-Klasse von Mengen. Solche Prozesse sind bereits im i.i.d.-Fall von NOLAN-POLLARD(1985) auf Glivenko-Cantelli-Verhalten hin behandelt worden. Ganz zentral für unser Vorgehen sind dabei die Lemmata 3.1.4 und 3.2.8.

Mit Hilfe von 3.1.4 sind wir in der Lage, eine bestmögliche stochastische Ungleichung herzuleiten. Das Korollar 3.2.2 liefert dabei die gleiche exponentielle Schranke wie in KIEFER(1961) (dort für die Gesamtheit aller Rechtecke im $\mathbb{R}^k$ anstelle von C) und verallgemeinert so ein entsprechendes Resultat aus HELMERS-JANSSEN-SERFLING(1985), die den U-Prozeß im i.i.d.-Fall und für eine spezielle, durch eine Kernfunktion h induzierte Vapnik-Chervonenkis-Klasse (totalgeordnet bzgl. " C "), untersuchen. Lemma 3.2.8 hingegen ermöglicht es uns, L_b -Aussagen unter relativ schwachen Bedingungen zu beweisen (die Halbmetrik wird dabei induziert durch die Varianzstruktur des Prozesses; vgl. hierzu Bedingung (C) in 3.3). Das Resultat 3.3.2 stellt dabei das Hauptergebnis von Kapitel III dar.

Ähnlich wie schon zu Ende von Kapitel II interessieren wir uns auch beim U-Prozeß für Aussagen im gewichteten Fall. Dabei beschränken wir uns jedoch auf die Situation $C = \{h^{-1}((-\infty, t]): t \in \mathbb{R}\}$, wobei h eine sogenannte Kernfunktion ist. Prozesse dieser Art wurden im i.i.d.-Fall bereits von SILVERMAN(1983) auf Verteilungskonvergenz hin untersucht. Das dortige Hauptresultat wird in Kapitel IV in zweierlei Hinsicht verallgemeinert: Zum einen arbeiten wir unter schwächeren Bedingungen an die Gewichtsfunktionen, und zum anderen fällt die i.i.d.-Voraussetzung weg. Darüberhinaus ist der Beweis relativ elementar – bei Silverman hingegen wird KMT (die Komlós-Major-Tusnády-Approximation) eingesetzt. Unsere Ergebnisse leiten wir selbstverständlich

wieder im Rahmen der L_b -Theorie her.

Weiterer Gegenstand der Betrachtungen in Kapitel IV sind Raten im Glivenko-Cantelli-Theorem für den i.i.d.-Fall, wobei wir entscheidend Gebrauch machen von geeigneten Martingaleigenschaften des U-Prozesses. Dabei stellt 4.4.2 für manche Kerne eine Verbesserung von Ergebnissen in NOLAN-POLLARD(1985) dar (allerdings werden dort auch Funktionenklassen als mögliche Parametermengen zugelassen). Den Abschluß der Arbeit bilden Betrachtungen darüber, ob die in 4.4 erhaltenen Ergebnisse optimal sind, wobei 4.4.7 zeigt, daß i.a. keine Verbesserung möglich ist. -

Was die Notation anbelangt, lehnen wir uns stets an GÄNSSLER-STUTE(1977) bzw. GÄNSSLER(1983) an. Im übrigen sei noch angemerkt, daß in allen Kapiteln die Folge der betrachteten Zufallselemente zwar über demselben Wahrscheinlichkeitsraum definiert ist, wir das jedoch für die möglichen "Häufungspunkte" nicht fordern (obwohl es o.E. möglich wäre durch eine geeignete Erweiterung des Wahrscheinlichkeitsraumes). Dies ist aber ohne Bedeutung, da der Begriff der L_b -Konvergenz lediglich von den Verteilungen der Zufallselemente abhängig ist.

Mein besonderer Dank gilt an dieser Stelle Herrn Prof. Dr. P. Gänßler für die Anleitung zu diesem Thema und seine fortwährende Unterstützung und Diskussionsbereitschaft.

KAPITEL I: EINE CHARAKTERISIERUNG DER L_b -KONVERGENZ VON ZUFALLSELEMENTEN IN FUNKTIONENRÄUMEN

1.1 Grundbegriffe und Definitionen

Sei (X, s) ein halbmétrischer Raum. Dann definieren wir:

1.1.1 Definition.

$$\ell^\infty(X) := \{f \in \mathbb{R}^X : f \text{ ist beschränkt}\}$$

$$S_0 := \{f \in \mathbb{R}^X : f \text{ ist gleichmäßig } s\text{-stetig und beschränkt}\}$$

$$d(f, g) := \sup\{|f(x) - g(x)| : x \in X\} \text{ für } f, g \in \ell^\infty(X)$$

$$\omega_f(\delta) := \sup\{|f(x) - f(y)| : x, y \in X \text{ mit } s(x, y) \leq \delta\} \text{ für } f \in \ell^\infty(X), \delta > 0$$

Weiter betrachten wir einen beliebigen Funktionenraum S , von dem wir lediglich voraussetzen wollen, daß gilt: $S_0 \subset S \subset \ell^\infty(X)$. Gegenstand der Untersuchungen im gesamten Verlauf des Kapitels I wird eine Folge $(X_n)_{n \in \mathbb{N}}$ von Zufallselementen über einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ mit Werten in S sein, die wir auf L_b -Konvergenz hin untersuchen wollen. Diesen Begriff der schwachen (Verteilungs-)Konvergenz findet man ausführlich behandelt in GÄNSSLER (1983), und unser Ziel wird es sein, auf der Basis der dortigen Ergebnisse eine Charakterisierung anhand des "Stetigkeitsmoduls" $\omega_f(\cdot)$ zu gewinnen. Aus diesem Grund haben wir uns auch in der Bezeichnungsweise mit (S_0, S, d) an diese Monographie angelehnt. Wir wollen noch einmal das für uns Wesentliche zusammenfassen:

$\mathcal{B}(S)$ bezeichne die Borel- σ -Algebra in (S, d) und es sei die

Meßbarkeitsforderung (M) erfüllt, d.h. es existiere eine Sub- σ -Algebra A von $\mathcal{B}(S)$ mit:

(M1) Für alle $g \in S$ ist $\phi_g: S \rightarrow \mathbb{R}$, $f \mapsto d(f, g)$ $A, \mathcal{B}^*$ -meßbar.

(M2) Für alle $n \in \mathbb{N}$ ist X_n, F, A -meßbar.

Dabei ist die Bedingung (M1) offensichtlich äquivalent zur Forderung, daß A die von den d -offenen Kugeln erzeugte σ -Algebra $\mathcal{B}_d(S)$ umfaßt, sodaß wir uns nun wegen (M2) in der Situation von Kapitel III in GÄNSSLER(1983) befinden (man setze für jedes $n \in \mathbb{N}$ $\mu_n := L(X_n)$); unter Berücksichtigung des dortigen Portmanteau-Theorems ((28), Seite 47) definieren wir:

1.1.2 Definition. $(X_n)_{n \in \mathbb{N}}$ heißt L_b -konvergent gegen ein $\tilde{F}, \mathcal{B}_b(S)$ -meßbares Zufallselement X_0 in S , definiert über $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ (in Zeichen: $X_n \xrightarrow{L_b} X_0$), falls erfüllt ist:

(i) $\int \phi \circ X_n d\mathbb{P} \rightarrow \int \phi \circ X_0 d\tilde{\mathbb{P}}$ für $n \rightarrow \infty$ und alle $\phi \in \mathcal{U}_b^d(S) := \{\phi' \in \mathbb{R}^S: \phi' \text{ ist gleichmäßig } d\text{-stetig, beschränkt und } \mathcal{B}_b(S), \mathcal{B}^*\text{-meßbar}\}.$

(ii) S_0 ist separabel und es gilt: $\tilde{\mathbb{P}}(X_0 \in S_0^C) = 1.$

(Man beachte, daß $S_0^C \in \mathcal{B}_b(S)$, falls S_0 separabel ist; dabei bezeichne S_0^C die abgeschlossene Hülle von S_0 in (S, d) .)

Weiter nennen wir $(X_n)_{n \in \mathbb{N}}$ relativ L_b -folgenkompakt, falls zu jeder Teilfolge von $(X_n)_{n \in \mathbb{N}}$ eine L_b -konvergente Teilteilfolge existiert.

Dann gilt als eine erste Charakterisierung (eine unmittelbare Folgerung aus dem Portmanteau-Theorem sowie Theorem 11a) in GÄNSSLER(1983)):

1.1.3 Theorem. Die beiden folgenden Aussagen sind äquivalent:

(i) $(X_n)_{n \in \mathbb{N}}$ ist relativ L_b -folgenkompakt.

(ii) Die Bedingungen B1) und B2) sind erfüllt:

B1) Für jede Folge $(\phi_m)_{m \in \mathbb{N}} \subset U_b^b(S)$ mit $\phi_m \rightarrow 0$ für $m \rightarrow \infty$

gilt: $\limsup_{n \rightarrow \infty} \int \phi_m \cdot X_n d\mathbb{P} \rightarrow 0$ für $m \rightarrow \infty$.

B2) S_0 ist separabel und für jedes $\phi \in U_b^b(S)$ mit $\phi \geq 1$ S_0^c

gilt: $\liminf_{n \rightarrow \infty} \int \phi \cdot X_n d\mathbb{P} \geq 1$.

Mit Theorem 1.1.3 als Ausgangspunkt wollen wir nun eine äquivalente Formulierung der relativen L_b -Folgenkompaktheit mittels des Stetigkeitsmoduls gewinnen. Für den Fall $X = [0,1]$ und $s(x,y) := |x-y|$ sowie $S = D([0,1])$ (definiert als der Raum der rechtsseitig stetigen Funktionen auf $[0,1]$ mit existierenden linksseitigen Limiten) ist dies bereits in Form von Theorem 1.17 in GÄNSSLER-HÄUSLER-SCHNEEMEIER(1984) enthalten. Es stellt sich nun heraus, daß der dortige Beweis sich im wesentlichen auf unsere sehr viel allgemeinere Situation übertragen läßt. Die dafür noch zusätzlich benötigten Hilfsmittel wollen wir nun bereitstellen.

1.2 Ein fundamentales Approximationslemma

Das folgende Lemma wird beim Beweis der angestrebten Charakterisierung sehr nützlich sein und überdies die Frage der Separabilität von S_0 klären.

1.2.1 Lemma. Falls (X, s) totalbeschränkt ist, existiert $S_1 \subset S_0$ mit folgender Eigenschaft: S_1 ist abzählbar und für jedes $f \in \ell^\infty(X)$ mit $\omega_f(\delta) \leq \varepsilon$ für $\delta > 0, \varepsilon > 0$ gibt es ein $g \in S_1$ mit $d(f, g) \leq 5\varepsilon$.

Beweis. Zu jedem $n \in \mathbb{N}$ existiert gemäß der vorausgesetzten Totalbeschränktheit von (X, s) ein E_n mit $|E_n| < \infty$ und der Eigenschaft: zu jedem $x \in X$ gibt es ein $e \in E_n$ mit $s(e, x) < n^{-1}$. Dann definieren wir für jedes $n \in \mathbb{N}$, $q \in \mathbb{Q}$ und $f \in \mathbb{Q}^{E_n}$

$$g(n, q, f)(x) := \sup\{f(e) - q \cdot s(e, x) : e \in E_n\}, \quad x \in X,$$

und betrachten die abzählbare Menge

$$S_1 := \{g(n, q, f) : n \in \mathbb{N}, q \in \mathbb{Q}, f \in \mathbb{Q}^{E_n}\}.$$

Wir werden nun zeigen, daß S_1 die im Lemma behaupteten Eigenschaften besitzt. Zunächst folgt aus der Totalbeschränktheit von (X, s) , daß für $x \in X$ gilt: $g_x \in S_0$ für $g_x(y) := s(x, y)$, $y \in X$, denn $\sup\{s(x, y) : y \in X\} \leq \sup\{2 + s(e, e') : e, e' \in E_1\} < \infty$. Damit ergibt sich aber unmittelbar, daß $S_1 \subset S_0$.

Es verbleibt also noch nachzuweisen, daß S_1 die geforderte Approximationseigenschaft besitzt. Hierzu fixieren wir ein $f \in \ell^\infty(X)$ mit $\omega_f(\delta) \leq \varepsilon$ für festes $\delta > 0$, $\varepsilon > 0$. Unser Ziel ist nun, ein Element aus S_1 zu finden, das f "gut genug" annähert. Dafür wählen wir ein $\delta' \in (0, \delta] \cap \mathbb{Q}$, ein $\varepsilon' \in [\varepsilon, \frac{5}{4}\varepsilon] \cap \mathbb{Q}$ sowie ein $M \in \{2 \cdot d(f, 0), \infty\} \cap \mathbb{Q}$ und betrachten die Gerade $K(t) := q \cdot t$, wobei $t \geq 0$ und $q := \max(\varepsilon', M) \cdot \delta'^{-1} \in \mathbb{Q}$.

Man erhält nun für alle $t \geq \delta^* := \varepsilon' \cdot \delta' \cdot \max(\varepsilon', M)^{-1}$:

$$(*) \quad \omega_f(t) \leq K(t)$$

Begründung zu (*): erfolgt durch Fallunterscheidung:

$\alpha)$ Sei $t \leq \delta'$. Dann ist $\omega_f(t) \leq \omega_f(\delta') \leq \epsilon' = q \cdot \delta^*$
 $\leq q \cdot t = K(t)$.

$\beta)$ Sei $t > \delta'$. Dann ist $\omega_f(t) \leq M \leq \max(M, \epsilon') = q \cdot \delta^*$
 $\leq q \cdot t = K(t)$.

Damit wäre also (*) gezeigt. Wir wählen nun ein $k \in \mathbb{N}$ mit $k^{-1} \leq \delta^*$ sowie ein $f^* \in \mathbb{Q}^{E_k}$ mit der Eigenschaft:

$\sup \{ |f^*(e) - f(e)| : e \in E_k \} \leq \frac{\epsilon'}{2}$ (f^* ist also eine "rationale Approximation" von f an den "Stützstellen" aus E_k). Für $g := g(k, q, f^*) \in S_1$ werden wir jetzt zeigen, daß $d(f, g) \leq 5\epsilon$. Zunächst einmal gilt für jedes $e \in E_k$:

$$\begin{aligned} 0 &\leq f^*(e) - f(e) - K(s(e, e)) + \frac{\epsilon'}{2} \\ &\leq \sup \{ f^*(e') - f(e) - K(s(e', e)) : e' \in E_k \} + \frac{\epsilon'}{2} \\ &= g(e) - f(e) + \frac{\epsilon'}{2} \\ &\leq \sup \{ \omega_f(s(e', e)) + f^*(e') - f(e') - K(s(e, e')) : e' \in E_k \} + \frac{\epsilon'}{2} \\ &\leq \sup \{ \omega_f(s(e', e)) - K(s(e', e)) : e' \in E_k \} + \epsilon' \\ &\leq 2\epsilon', \\ &\text{denn falls } s(e', e) \leq \delta^*, \text{ folgt } \omega_f(s(e', e)) \leq \omega_f(\delta^*) \leq \epsilon'; \\ &\text{ist hingegen } s(e', e) > \delta^*, \text{ folgt wegen } (*): \omega_f(s(e', e)) \leq \\ &\quad K(s(e', e)). \end{aligned}$$

Insgesamt erhalten wir damit:

$$(**) \sup \{ |g(e) - f(e)| : e \in E_k \} \leq 2\epsilon'.$$

Weiter gilt für beliebige $x, x' \in X$:

$$\begin{aligned} g(x) - g(x') &= \sup \{ f^*(e) - K(s(x, e)) : e \in E_k \} - \\ &\quad \sup \{ f^*(e') - K(s(x', e')) : e' \in E_k \} \\ &\leq \sup \{ f^*(e) - K(s(x, e)) - f^*(e) + K(s(x', e)) : e \in E_k \} \end{aligned}$$

$$\begin{aligned} &\leq \sup \{ K(s(x, x') + s(x, e)) - K(s(x, e)) : e \in E_k \} \\ &= K(s(x, x')) , \end{aligned}$$

also insbesondere für den Fall $s(x, x') \leq \delta^*$:

$$|g(x) - g(x')| \leq K(s(x, x')) \leq q \cdot \delta^* = \epsilon'.$$

Dies liefert aber unmittelbar die Gültigkeit von

$$(***) \quad \omega_g(\delta^*) \leq \epsilon'.$$

Aus (**) und (***) ergibt sich nun sofort die Behauptung des Lemmas, denn zu jedem $x \in X$ existiert ein $e = e(x) \in E_k$ mit $s(x, e) < k^{-1} \leq \delta^*$ und es folgt:

$$|g(x) - f(x)| \leq |g(x) - g(e)| + |g(e) - f(e)| + |f(e) - f(x)|$$

$$\begin{aligned} &(**) \\ &\leq \omega_g(\delta^*) + 2\epsilon' + \omega_f(\delta^*) \end{aligned}$$

$$\begin{aligned} &(***) \\ &\leq 4\epsilon' \leq 5\epsilon. \end{aligned}$$

□

1.2.2 Bemerkung. Das Approximationslemma 1.2.1 wurde in GÄNSSLER-HÄUSLER-SCHNEEMEIER(1984) in der speziellen Situation $X = [0, 1]$, $s(x, y) = |x - y|$ "vertreten" durch ein Approximieren mit Hilfe von Polygonzügen. Dort benutzte man für jedes $f \in \mathbb{R}^X$ mit $\omega_f(\delta) \leq \epsilon$ für $\delta, \epsilon > 0$ als geeignetes g :
Es existiert ein $N \in \mathbb{N}$ mit $N^{-1} \leq \delta$. Setze nun $g(\frac{i}{N}) := f(\frac{i}{N})$ für $i = 0, 1, \dots, N$ und g auf dem Intervall $[\frac{i-1}{N}, \frac{i}{N}]$ linear. Dann ist $g \in S_0$ und $d(f, g) \leq 2\epsilon$.

Als unmittelbare Folgerung aus Lemma 1.2.1 erhalten wir weiterhin (in gewisser Weise eine Umkehrung):

1.2.3 Lemma. (S_0, d) ist genau dann separabel, falls der Raum (X, s) totalbeschränkt ist.

Beweis. Falls (X, s) totalbeschränkt ist, folgt sofort gemäß Lemma 1.2.1 die Separabilität von S_0 (mit S_1 als geeignetem Separator).

Es verbleibt nun also nur mehr, die umgekehrte Implikation nachzuweisen. Den Beweis hierzu führen wir indirekt. Sei also (S_0, d) separabel und wir nehmen an: (X, s) ist nicht totalbeschränkt. Dann findet man ein $\delta > 0$ und (rekursiv definiert) eine Folge $(e_n)_{n \in \mathbb{N}} \subset X$ mit $s(e_i, e_j) \geq \delta$ für alle $i \neq j$, wobei $i, j \in \mathbb{N}$.

Für jedes $E \subset \mathbb{N}$ setzen wir

$$f_E(x) := \sum_{n \in E} \max\left(\frac{\delta}{3} - s(x, e_n), 0\right) \quad \text{für } x \in X.$$

Aus der Dreiecksungleichung ergibt sich nun sofort für jedes $x \in X$, daß höchstens ein $n \in \mathbb{N}$ existiert mit $s(x, e_n) < \frac{\delta}{3}$, also folgt unmittelbar: $d(f_E, 0) \leq \frac{\delta}{3}$ und daher wegen obiger Definition $f_E \in S_0$.

Andererseits ist jedoch $d(f_E, f_{E'}) = \frac{\delta}{3}$ für $E, E' \subset \mathbb{N}$ mit $E \neq E'$, da $f_E(e_n) = \frac{\delta}{3} \cdot 1_E(n)$ für alle $n \in \mathbb{N}$ und $E \subset \mathbb{N}$. Für jede dichte Teilmenge S' von S_0 müßte also gelten $\text{card}(S') \geq \text{card}(\mathcal{P}(\mathbb{N}))$, d.h. (S_0, d) wäre nicht mehr separabel (Widerspruch). $\square$

Nach diesen grundsätzlichen Überlegungen wollen wir uns wieder der Untersuchung der L_b -Konvergenz zuwenden.

1.3 Charakterisierung der relativen L_b -Folgenkompaktheit anhand des Stetigkeitsmoduls

Wie bereits in Teil 1.1 erwähnt, wollen wir in diesem Abschnitt eine geeignete Modifikation der Beweise in GÄNSSLER-HÄUSLER-SCHNEEMEIER(1984) vornehmen (natürlich mit Hilfe der Ergebnisse aus 1.2). Hierzu betrachten wir wieder die Folge $(X_n)_{n \in \mathbb{N}}$ aus 1.1 und setzen voraus:

(V1) (X, s) ist ein totalbeschränkter halbmétrischer Raum.

(V2) Die Meßbarkeitsforderung (M) aus 1.1 möge erfüllt sein.

Die Forderung (V1) liefert uns gemäß Lemma 1.2.3 nun unmittelbar die Separabilität von S_0 , so daß wir diesen Aspekt der L_b -Konvergenz (vgl. Definition 1.1.2) im weiteren Verlauf nicht mehr explizit erwähnen werden.

Wir wollen zunächst noch zwei Ergebnisse aus Kapitel 3 in GÄNSSLER(1983) hier mit aufführen, die für den Beweis des Charakterisierungstheorems benötigt werden:

1.3.1 Theorem (Teil des Portmanteau-Theorems). Für jedes $\tilde{F}, B_b(S)$ -meßbare X_0 in S über $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ sind die beiden folgenden Aussagen äquivalent, sofern $\tilde{\mathbb{P}}(X_0 \in S_0) = 1$:

(i) Für alle $\phi \in \mathcal{U}_b^b(S)$ gilt: $\lim_{n \rightarrow \infty} \int \phi \circ X_n d\mathbb{P} = \int \phi \circ X_0 d\tilde{\mathbb{P}}$.

(ii) Für alle d -abgeschlossenen Teilmengen F von S gilt:

$$\lim_{n \rightarrow \infty} \sup L(X_n)^*(F) \leq \tilde{\mathbb{P}}(X_0 \in F \cap S_0).$$

1.3.2 Bemerkung. In 1.3.1 wie auch im folgenden bezeichne $L(X_n)^*$ stets das zu $L(X_n)|A$ gehörige äußere Maß (ebenso sei $\mathbb{P}^*$ das zu $\mathbb{P}|F$ gehörige äußere Maß). Im übrigen wurde in der Formulierung des Theorems Gebrauch gemacht von der Tatsache, daß S_0 abgeschlossen ist in S , also insbesondere gilt: $S_0 \cap F \in \mathcal{B}_b(S)$ für alle d -abgeschlossenen Teilmengen F von S (mithin ist auch $S_0 \in \mathcal{B}_b(S)$).

1.3.3 Theorem (Stetigkeitssatz). Wir setzen voraus, daß ein $\tilde{F}, \mathcal{B}_b(S)$ -meßbares X_0 in S über $(\tilde{\Omega}, \tilde{F}, \tilde{\mathbb{P}})$ existiert mit $X_n \xrightarrow{L_b} X_0$. Dann folgt für jede d -stetige und $A, \mathcal{B}_k^*$ -meßbare Funktion g von S nach $\mathbb{R}^k$: $g \circ X_n \xrightarrow{L} (g \cdot 1_{S_0}) \circ X_0$.

Nach diesen grundsätzlichen Bemerkungen sind wir nun in der Lage, die gewünschte Charakterisierung herzuleiten.

1.3.4 Theorem. Die folgenden vier Aussagen sind äquivalent:

- (i) $(X_n)_{n \in \mathbb{N}}$ erfüllt die Bedingungen B1) und B2) aus Theorem 1.1.3.
- (ii) $(X_n)_{n \in \mathbb{N}}$ ist relativ L_b -folgenkompakt.
- (iii) Die Bedingungen (A) und (B) sind erfüllt, d.h. es gilt:
 - (A) $\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}^*(\omega_{X_n}(\delta) \geq \varepsilon) = 0$ für alle $\varepsilon > 0$.
 - (B) $\lim_{K \uparrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}^*(|X_n(x)| \geq K) = 0$ für alle $x \in X$.
- (iv) Die Bedingung (A) aus (iii) sei erfüllt sowie:
 - (B') $\lim_{K \uparrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(d(X_n, 0) \geq K) = 0$.

Entsprechend zu GÄNSSLER-HÄUSLER-SCHNEEMEIER(1984) basiert der Beweis zu Theorem 1.3.4 auf den beiden folgenden Lemmata:

1.3.5 Lemma. Sei $(K_n)_{n \in \mathbb{N}}$ eine monoton fallende Folge von Teilmengen von S mit der folgenden Eigenschaft (K):

(K) Jede Folge $(f_n)_{n \in \mathbb{N}}$ mit $f_n \in K_n$ für $n \in \mathbb{N}$ besitzt einen Häufungspunkt in (S, d) .

Dann gilt für jede Folge $(\phi_n)_{n \in \mathbb{N}} \subset C(S) := \{\phi \in \mathbb{R}^S : \phi \text{ ist } d\text{-stetig}\}$ mit $\phi_n \rightarrow 0$ für $n \rightarrow \infty$: $\sup\{\phi_n(f) : f \in K_n\} \rightarrow 0$ für $n \rightarrow \infty$.

1.3.6 Lemma. Seien $(\epsilon_k)_{k \in \mathbb{N}} \subset (0, \infty)$ und $(\delta_k)_{k \in \mathbb{N}} \subset (0, \infty)$ mit $\epsilon_k \rightarrow 0$ für $k \rightarrow \infty$ sowie $K > 0$ fest vorgegeben. Dann gilt für jede Folge $(f_n)_{n \in \mathbb{N}}$ mit $f_n \in \bigcap_{i=1}^n \{f \in S : \omega_f(\delta_i) < \epsilon_i \text{ und } d(f, 0) < K\} =: K_n$ für $n \in \mathbb{N}$:

$(f_n)_{n \in \mathbb{N}}$ besitzt einen Häufungspunkt in (S_0, d) , d.h. insbesondere erfüllt $(K_n)_{n \in \mathbb{N}}$ die Bedingung (K) aus Lemma 1.3.5.

Beweis zu 1.3.5. Im Gegensatz zu GÄNSSLER-HÄUSLER-SCHNEEMEIER(1984) wird hier ein direkter Beweis erbracht, der jedoch dieselbe Idee benutzt. Sei also eine Folge $(\phi_n)_{n \in \mathbb{N}} \subset C(S)$ mit $\phi_n \rightarrow 0$ für $n \rightarrow \infty$ sowie ein beliebiges $\epsilon > 0$ fest vorgegeben. Dann existiert zu jedem $n \in \mathbb{N}$ ein $f_n \in K_n$ mit

$$(*) \sup\{\phi_n(f) : f \in K_n\} \leq \phi_n(f_n) + \epsilon \text{ (wobei o.E. } \phi_1 \leq 1 \text{)}.$$

Des weiteren existiert gemäß Voraussetzung (K) ein $f \in S$ sowie eine Teilfolge $(f_{n_k})_{k \in \mathbb{N}}$ von $(f_n)_{n \in \mathbb{N}}$ mit $d(f_{n_k}, f) \rightarrow 0$ für $k \rightarrow \infty$. Wegen der vorausgesetzten punktwweisen Monotonie von $(\phi_n)_{n \in \mathbb{N}}$ und der Monotonie von $(K_n)_{n \in \mathbb{N}}$ folgt sofort für alle $m \in \mathbb{N}$ (da ϕ_m d -stetig ist):

$$\begin{aligned}
 \lim_{n \rightarrow \infty} \sup \{ \phi_n(f) : f \in K_n \} &\stackrel{(*)}{\leq} \lim_{k \rightarrow \infty} \sup \phi_{n_k}(f_{n_k}) + \varepsilon \\
 &\leq \lim_{k \rightarrow \infty} \sup \phi_m(f_{n_k}) + \varepsilon \\
 &= \phi_m(f) + \varepsilon
 \end{aligned}$$

Mit $m \rightarrow \infty$ folgt nun sofort die Behauptung, da ε beliebig gewählt war. □

Beweis zu 1.3.6. Gemäß (V1) existiert eine abzählbare Teilmenge A von X mit $A^C = X$. Da außerdem $[-K, K]$ kompakt ist, existiert (mittels Diagonalverfahren) eine Teilfolge $(f_{n_k})_{k \in \mathbb{N}}$ von $(f_n)_{n \in \mathbb{N}}$ (also mit $n_1 < n_2 < \dots$) so, daß für jedes $x \in A$ die Folge $(f_{n_k}(x))_{k \in \mathbb{N}}$ konvergent ist. Wir definieren nun für jedes $x \in A$ $g(x) := \lim_{k \rightarrow \infty} f_{n_k}(x)$ und zeigen zunächst:

(*) g ist gleichmäßig stetig auf A .

Begründung für (*): Zu beliebigem $\varepsilon > 0$ existiert ein

$k_0 \in \mathbb{N}$ mit $\varepsilon_{n_{k_0}} \leq \varepsilon$. Dann gilt für $a, b \in A$ mit

$s(a, b) \leq \delta_{n_{k_0}}$ und alle $k \geq k_0$:

$$\begin{aligned}
 |g(a) - g(b)| &\leq |g(a) - f_{n_k}(a)| + |f_{n_k}(a) - f_{n_k}(b)| + \\
 &\quad |f_{n_k}(b) - g(b)| \\
 &\leq \omega_{f_{n_k}}(\delta_{n_{k_0}}) + 2 \sup \{ |f_{n_k}(x) - g(x)| : x = a, b \} \\
 &\leq \varepsilon + 2 \sup \{ |f_{n_k}(x) - g(x)| : x = a, b \},
 \end{aligned}$$

d.h. mit $k \rightarrow \infty$ folgt $|g(a) - g(b)| \leq \varepsilon$.

Somit ist (*) gezeigt, und wir erhalten daraus unmittelbar

die Existenz einer gleichmäßig stetigen Fortsetzung $f: X \rightarrow \mathbb{R}$ von g (die natürlich eindeutig bestimmt ist). Wir zeigen nun:

(**) $d(f_{n_k}, f) \rightarrow 0$ für $k \rightarrow \infty$, also ist insbesondere f ein Häufungspunkt von $(f_n)_{n \in \mathbb{N}}$ in (S_0, d) und Lemma 1.3.6 damit bewiesen.

Begründung für (**): Wir betrachten wieder ein beliebiges $\varepsilon > 0$ und wählen dazu ein $k_0 \in \mathbb{N}$ mit $\varepsilon_{n_{k_0}} \leq \frac{\varepsilon}{2}$ sowie

$\omega_f(\delta_{n_{k_0}}) \leq \frac{\varepsilon}{2}$ (zur Existenz vgl. Begründung zu (*)). Außerdem gibt es gemäß (V1) eine endliche Teilmenge E von A

mit der Eigenschaft: Für jedes $x \in X$ existiert ein $e_x \in E$ mit $s(x, e_x) \leq \delta_{n_{k_0}}$, d.h. es folgt für alle $k \geq k_0$:

$$\begin{aligned} |f_{n_k}(x) - f(x)| &\leq |f_{n_k}(x) - f_{n_k}(e_x)| + |f_{n_k}(e_x) - f(e_x)| + \\ &\quad |f(e_x) - f(x)| \\ &\leq \varepsilon + \sup\{|f_{n_k}(e) - f(e)| : e \in E\}. \end{aligned}$$

Insgesamt folgt also mit $k \rightarrow \infty$:

$\limsup_{k \rightarrow \infty} d(f_{n_k}, f) \leq \varepsilon$. Da ε beliebig gewählt war, folgt

(**) und damit die Behauptung. □

Beweis zu Theorem 1.3.4. Wir werden die folgende Implikationskette beweisen: (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Rightarrow$ (i). Dabei ist (i) $\Rightarrow$ (ii) eine unmittelbare Konsequenz aus Theorem 1.1.3.

(ii) $\Rightarrow$ (iii):

Wir zeigen lediglich, daß (A) erfüllt ist ((B) folgt analog).

Wie in GÄNSSLER-HÄUSLER-SCHNEEMEIER(1984) erfolgt der Beweis indirekt. Wir nehmen an, daß ein $\varepsilon > 0$ und $\eta > 0$ existieren mit

$\lim_{n \rightarrow \infty} \sup \mathbb{P}^* (\omega_{X_n} (\frac{1}{N}) \geq \varepsilon) > \eta$ für alle $N \in \mathbb{N}$.

Dann existiert eine Teilfolge $(n_N)_{N \in \mathbb{N}}$ von $\mathbb{N}$ mit

(+) $\mathbb{P}^* (\omega_{X_{n_N}} (\frac{1}{N}) \geq \varepsilon) > \eta$ für alle $N \in \mathbb{N}$.

Wegen der vorausgesetzten relativen L_b -Folgenkompaktheit von

$(X_n)_{n \in \mathbb{N}}$ existiert jedoch eine Teilfolge $(X_{n_{N_k}})_{k \in \mathbb{N}}$ von

$(X_{n_N})_{N \in \mathbb{N}}$, die L_b -konvergent ist gegen ein $\tilde{F}, \mathcal{B}_b(S)$ -meßbares X_0 in S über $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ mit $\tilde{\mathbb{P}}(X_0 \in S_0) = 1$. Mittels 1.3.1 folgt nun (man beachte, daß $\{f \in S: \omega_f(\delta) \geq \varepsilon\}$ für jedes $\delta > 0$ abgeschlossen ist in (S, d)):

(++) $\lim_{k \rightarrow \infty} \sup \mathbb{P}^* (X_{n_{N_k}} \in \{f \in S: \omega_f(\delta) \geq \varepsilon\}) \leq \tilde{\mathbb{P}}(\{\omega_{X_0}(\delta) \geq \varepsilon\} \cap \{X_0 \in S_0\}) \rightarrow 0$ für $\delta > 0$.

Aus (+) folgt aber für jedes $\delta > 0$:

$\lim_{k \rightarrow \infty} \inf \mathbb{P}^* (X_{n_{N_k}} \in \{f \in S: \omega_f(\delta) \geq \varepsilon\}) \geq \eta$ im Widerspruch zu (++) .

(iii) $\Rightarrow$ (iv):

Für jedes $\delta > 0$ bezeichne E_δ eine endliche Teilmenge mit der Eigenschaft: Für alle $x \in X$ existiert ein $e \in E_\delta$ mit $s(x, e) \leq \delta$ (vgl. hierzu (V1)). Dann erhalten wir für $\varepsilon > 0, \delta > 0$ sowie $K > 0$:

$\{f \in S: \omega_f(\delta) < \varepsilon\} \cap \{f \in S: |f(e)| < K \text{ für } e \in E_\delta\} \subset \{f \in S: d(f, 0) < K + \varepsilon\}$,

d.h. es folgt wegen (A) und (B) für alle $\delta > 0$:

$\lim_{K \rightarrow \infty} \lim_{n \rightarrow \infty} \sup \mathbb{P}(d(X_n, 0) \geq K) \leq \lim_{K \rightarrow \infty} \lim_{n \rightarrow \infty} \sup [\mathbb{P}^*(\omega_{X_n}(\delta) \geq 1) + \sum_{e \in E_\delta} \mathbb{P}^*(|X_n(e)| \geq K-1)]$
 (B)
 $= \lim_{n \rightarrow \infty} \sup \mathbb{P}^*(\omega_{X_n}(\delta) \geq 1) + 0$ für $\delta > 0$.

Damit wäre aber (B') gezeigt. (Da (B') trivialerweise die Bedingung (B) impliziert, sind (iii) und (iv) äquivalent.)
(iv) $\Rightarrow$ (i):

Hier lehnen wir uns wieder an GÄNSSLER-HÄUSLER-SCHNEEMEIER(1984) an und zeigen:

Teil 1: Für beliebiges $\eta > 0$ existieren $(\varepsilon_k)_{k \in \mathbb{N}}$ und $(\delta_k)_{k \in \mathbb{N}}$

sowie ein $K > 0$ wie in 1.3.6 so, daß für alle $n \geq n_1(\eta)$

$$\text{gilt: } \mathbb{P}^* \left(\bigcup_{k=1}^n \{ \omega_{X_n}(\delta_k) \geq \varepsilon_k \} \cup \{ d(X_n, O) \geq K \} \right) \leq \eta.$$

Beweis zu Teil 1:

Sei $\eta > 0$ beliebig. Sei weiter eine Folge $(\tilde{\varepsilon}_k)_{k \in \mathbb{N}}$ mit $\tilde{\varepsilon}_k \downarrow 0$ für $k \rightarrow \infty$ fest vorgegeben. Wegen der Gültigkeit von (A) existiert zu jedem $k \in \mathbb{N}$ ein $\tilde{\delta}_k > 0$ mit $\limsup_{n \rightarrow \infty} \mathbb{P}^*(\omega_{X_n}(\tilde{\delta}_k) \geq \tilde{\varepsilon}_k) \leq \frac{\eta}{4} \cdot 2^{-k}$.

Somit folgt für jedes $N \in \mathbb{N}$: $\limsup_{n \rightarrow \infty} \mathbb{P}^* \left(\bigcup_{k=1}^N \{ \omega_{X_n}(\tilde{\delta}_k) \geq \tilde{\varepsilon}_k \} \right) \leq \frac{\eta}{4}$,

d.h. wir erhalten zu $N \in \mathbb{N}$ ein $n_N \in \mathbb{N}$ mit $n_N > n_{N-1}$ ($n_0 := 0$) so,

$$\text{daß } \mathbb{P}^* \left(\bigcup_{k=1}^N \{ \omega_{X_{n_N}}(\tilde{\delta}_k) \geq \tilde{\varepsilon}_k \} \right) \leq \frac{\eta}{2} \text{ für alle } n \geq n_N \quad (*).$$

Wir setzen nun: $\varepsilon_k := \tilde{\varepsilon}_1$, $\delta_k := \tilde{\delta}_1$ für $k=1, \dots, n_1-1$.

$$\varepsilon_k := \tilde{\varepsilon}_N, \quad \delta_k := \tilde{\delta}_N \text{ für } k=n_N, \dots, n_{N+1}-1 \quad (N \in \mathbb{N}).$$

Dann gilt für jedes $n \geq n_1$: es existiert genau ein $N \in \mathbb{N}$ mit $n \in \{n_N, \dots, n_{N+1}-1\}$ und somit folgt:

$$\mathbb{P}^* \left(\bigcup_{k=1}^n \{ \omega_{X_n}(\delta_k) \geq \varepsilon_k \} \right) = \mathbb{P}^* \left(\bigcup_{k=n_1}^{n_{N+1}-1} \{ \omega_{X_n}(\delta_k) \geq \varepsilon_k \} \right)$$

$$\stackrel{(*)}{=} \mathbb{P}^* \left(\bigcup_{i=1}^N \{ \omega_{X_{n_N}}(\tilde{\delta}_i) \geq \tilde{\varepsilon}_i \} \right) \leq \frac{\eta}{2} \quad (\text{wegen } n \geq n_N).$$

Zusammen mit (B') folgt nun sofort die Behauptung von Teil 1.

Teil 2: Hier werden wir nun entscheidend Gebrauch machen von Lemma 1.2.1 (in Verbindung mit 1.3.5 bzw. 1.3.6).

Zunächst zeigen wir, daß die Bedingung B1) erfüllt ist.

Sei dazu $(\phi_m)_{m \in \mathbb{N}} \subset U_b^b(S)$ mit $\phi_m \rightarrow 0$ für $m \rightarrow \infty$ sowie $\phi_1 \leq 1$ (o.E.).

Dann existiert zu beliebigem $\eta > 0$ gemäß Teil 1 und Lemma

1.3.6 eine monoton wachsende Folge $(K_n)_{n \in \mathbb{N}}$ von d-abgeschlossenen Teilmengen von S mit den Eigenschaften:

$(K_n)_{n \in \mathbb{N}}$ erfüllt (K) aus 1.3.5 und $\mathbb{P}^*(X_n \in K_n) \leq \frac{\eta}{2}$ für alle $n \geq n_1$.

Weiter existiert gemäß 1.3.5 ein $m_1 \in \mathbb{N}$ mit $m_1 \geq n_1$ so, daß

$\sup\{\phi_m(f) : f \in K_m\} \leq \frac{\eta}{2}$ für alle $m \geq m_1$. Dann gilt für alle $m, n \in \mathbb{N}$ mit $n \geq m \geq m_1$:

$$\begin{aligned} \int \phi_m \circ X_n d\mathbb{P} &= \int [(\phi_m \circ X_n) \cdot 1_{\{X_n \in K_m\}} + (\phi_m \circ X_n) \cdot 1_{\{X_n \in K_m^c\}}] d\mathbb{P} \\ &\leq \frac{\eta}{2} + \mathbb{P}^*(X_n \in K_m) \\ &\leq \frac{\eta}{2} + \mathbb{P}^*(X_n \in K_n) \leq \eta, \end{aligned}$$

also insgesamt $\limsup_{n \rightarrow \infty} \int \phi_m \circ X_n d\mathbb{P} \leq \eta$ für alle $m \geq m_1$.

Da $\eta > 0$ beliebig gewählt war, folgt hieraus unmittelbar die Gültigkeit von B1). Wir wollen uns nun dem Beweis von B2) zuwenden.

Dazu betrachten wir ein beliebiges $\phi \in U_b^b(S)$ mit $\phi \geq 1_{S_0}$; sei darüberhinaus $\eta > 0$ wieder beliebig gewählt. Wegen der vorausgesetzten gleichmäßigen Stetigkeit von ϕ existiert dann ein $\delta = \delta(\eta) > 0$ mit $\phi \geq (1-\eta) \cdot 1_{S_0^\delta}$,

wobei $S_0^\delta := \{f \in S : d(f, g) < \delta \text{ für ein } g \in S_0\} \in \mathcal{B}_b(S)$ (vgl. hierzu GÄNSSLER(1983), Lemma 11).

Weiter existieren (vgl. Teil 1) Folgen $(\varepsilon_k)_{k \in \mathbb{N}}$ und $(\delta_k)_{k \in \mathbb{N}}$ $\subset (0, \infty)$ mit $\varepsilon_k \downarrow 0$ für $k \rightarrow \infty$ so, daß

$$\mathbb{P}^*(\omega_{X_n}(\delta_n) \geq \varepsilon_n) \leq \eta \quad \text{für alle } n \geq n_1 = n_1(\eta).$$

Gemäß Lemma 1.2.1 ergibt sich hieraus unmittelbar:

$$\mathbb{P}(X_n \in \bigcup_{O} S_O^{5\varepsilon_n}) \leq \eta \quad \text{für alle } n \geq n_1.$$

Weiter existiert ein $n_2 \geq n_1$ mit $5\varepsilon_{n_2} \leq \delta$ und wir erhalten somit für alle $n \geq n_2$:

$$\begin{aligned} \int \phi \circ X_n d\mathbb{P} &= \int (\phi \circ X_n) \cdot 1_{\{X_n \in \bigcup_{O} S_O^{5\varepsilon_n}\}} d\mathbb{P} + \int (\phi \circ X_n) \cdot 1_{\{X_n \in S_O^{5\varepsilon_n}\}} d\mathbb{P} \\ &\geq (1-\eta) \cdot \mathbb{P}(X_n \in S_O^{5\varepsilon_n}) \\ &\geq (1-\eta)^2. \end{aligned}$$

Damit ist aber B2) gezeigt und der Beweis zu Theorem 1.3.4 beendet.

□

1.3.7 Bemerkung.

(i) Die Bedingung (B) in 1.3.4 wird bei unseren Anwendungen stets automatisch erfüllt sein, denn es gilt z.B. im Falle von empirischen C-Prozessen $(X_n)_{n \in \mathbb{N}}$ vermöge der Tschebyscheff-Ungleichung für jedes $n \in \mathbb{N}$:

$$\mathbb{P}(|X_n(C)| \geq K) \leq (2K)^{-2} \quad \text{für alle } n \in \mathbb{N}, K > 0 \text{ und Werte } C \text{ aus dem Parameterraum (in diesem Fall ein Mengensystem).}$$

(ii) Im Hinblick auf die Aussagen in Theorem 1.3.4 bemerken wir: Sowohl in (iii) wie auch (iv) kann durchwegs an Stelle von $\mathbb{P}^*(X_n \in \cdot)$ auch $L(X_n)^*(\cdot)$ gesetzt werden (wobei $L(X_n)$

die durch X_n auf A induzierte Wahrscheinlichkeitsverteilung bezeichne). Dazu beachte man, daß in 1.3.4 (iv) äquivalent ist zu (iv') $\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}^*(X_n \in K(\delta, \epsilon)) = 0$ für alle $\epsilon > 0$,

$$\text{wobei } K(\delta, \epsilon) := \{f \in S: \omega_f(\delta) > \epsilon \text{ oder } d(f, 0) > \delta^{-1}\}.$$

Für rationales $\epsilon, \delta \in (0, 1]$ folgt nun (vgl. hierzu den Beweis von Lemma 1.2.1):

Für jedes $f \in K(\delta, \epsilon)$ existiert ein $g \in S_0$ mit den Eigenschaften $d(g, f) < 5\epsilon$ sowie $\omega_g(\delta^*) \leq \epsilon$, wobei $\delta^* \geq \frac{\epsilon}{2} \delta^2$. Somit gilt aber auch $\omega_g(\frac{\epsilon}{2} \delta^2) \leq \epsilon$ und $d(g, 0) \leq \delta^{-1} + 5\epsilon$, also insgesamt (wegen $\delta^{-1} + 5\epsilon \leq \frac{6}{\epsilon} \delta^{-2}$ und $\frac{6}{\epsilon} \delta^{-2} + 5\epsilon \leq \frac{11}{\epsilon} \delta^{-2}$):

$$(+)\ C(\delta, \epsilon) \subset \left(K\left(\frac{\epsilon}{6} \delta^2, \epsilon\right) \cap S_0 \right)^{5\epsilon} =: C(\delta, \epsilon) \subset C\left(\frac{\epsilon}{11} \delta^2, 11\epsilon\right).$$

Wegen $C(\delta, \epsilon) \in \mathcal{B}_b(S)$ (vgl. GÄNSSLER(1983), Lemma 11) folgt für $n \in \mathbb{N}$:

$$\begin{aligned} L(X_n)^* \left(K\left(\frac{\epsilon}{11} \delta^2, 11\epsilon\right) \right) &\stackrel{(+)}{\leq} L(X_n)(C(\delta, \epsilon)) \\ &= \mathbb{P}(X_n \in C(\delta, \epsilon)) \\ &\leq \mathbb{P}^*(X_n \in K(\delta, \epsilon)). \end{aligned}$$

Da weiterhin gemäß der Definition des äußeren Maßes unmittelbar gilt: $\mathbb{P}^*(X_n \in K(\delta, \epsilon)) \leq L(X_n)^*(K(\delta, \epsilon))$, erhalten wir sofort die Äquivalenz von (iv') mit

$$(iv'') \quad \lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} L(X_n)^*(K(\delta, \epsilon)) = 0 \text{ für alle } \epsilon > 0.$$

Andererseits ist jedoch (iv'') wiederum äquivalent zur Aussage (iv''') Die Bedingungen (C) und (D') sind erfüllt, d.h:

$$(C) \quad \lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} L(X_n)^* (\{f \in S: \omega_f(\delta) \geq \epsilon\}) = 0 \text{ für } \epsilon > 0.$$

$$(D') \quad \lim_{K \uparrow \infty} \limsup_{n \rightarrow \infty} L(X_n) (\{f \in S: d(f, 0) \geq K\}) = 0.$$

Ganz offensichtlich ist aber auch (iv''') äquivalent zu:

(iii') Es gilt (C) aus (iv''') sowie

$$(D) \lim_{K \uparrow \infty} \limsup_{n \rightarrow \infty} L(X_n)^* (\{f \in S: |X_n(x)| \geq K\}) = 0 \text{ für } x \in X.$$

Insgesamt erhalten wir damit aus diesen Überlegungen, daß sich Theorem 1.3.4 wie folgt formulieren läßt:

(i) $\Leftrightarrow$ (ii) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iii') $\Leftrightarrow$ (iv) $\Leftrightarrow$ (iv''').

1.4 Charakterisierung der L_b -Konvergenz und hinreichende

Kriterien

Wir wollen nun die Ergebnisse von 1.3 benützen, um eine Charakterisierung der L_b -Konvergenz herzuleiten. Es ist klar, daß hierfür mit Theorem 1.3.4 bereits die Hauptarbeit geleistet wurde. Was noch zu tun verbleibt, ist lediglich, eine Identifizierung aller möglichen Limiten einer Folge $(X_n)_{n \in \mathbb{N}}$ sicherzustellen (vgl. GÄNSSLER-HÄUSLER-SCHNEEMEIER(1984)). Dies muß natürlich anhand der Konvergenz der endlichdimensionalen Randverteilungen der $X_n, n \in \mathbb{N}$, erfolgen. Hierzu müssen wir jedoch die bisherigen Meßbarkeitsvoraussetzungen noch etwas verschärfen.

Wie im vorangehenden Abschnitt fordern wir für eine Folge $(X_n)_{n \in \mathbb{N}}$ von Zufallselementen in S die Gültigkeit von (V1) sowie (V2) und verlangen zusätzlich

(M3) Für alle $x \in X$ ist $\pi_x(S): S \rightarrow \mathbb{R}$, definiert durch

$$\pi_x(S)(f) := f(x), A, \mathcal{B}^* \text{-meßbar.}$$

(M2) zusammen mit (M3) impliziert, daß für alle $x \in X$ $X_n(x)$ eine zufällige Variable (also $F, \mathcal{B}^*$ -meßbar) ist, wobei letzteres

nichts anderes bedeutet als $F, \sigma(\{\pi_x(S) : x \in X\})$ -Meßbarkeit von X_n (welche im Fall von empirischen C-Prozessen immer erfüllt ist). Unter (V1), (V2) und (M3) gilt nun das folgende Theorem (vgl. Theorem 1.11 in GÄNSSLER-HÄUSLER-SCHNEE-MEIER(1984)):

1.4.1 Theorem. Die beiden folgenden Aussagen sind äquivalent:

- (i) Es existiert ein $\tilde{F}, \mathcal{B}_b(S)$ -meßbares X_0 (über einem $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$) in S mit $X_n \xrightarrow{L_b} X_0$.
- (ii) $(X_n)_{n \in \mathbb{N}}$ ist relativ L_b -folgenkompakt und die endlich-dimensionalen Randverteilungen von $(X_n)_{n \in \mathbb{N}}$ konvergieren in Verteilung, d.h. für alle $k \in \mathbb{N}$ und $(x_1, \dots, x_k) \in X^k$ existiert ein $\mathcal{B}_k^*$ -meßbarer Zufallsvektor Y mit $(X_n(x_1), \dots, X_n(x_k)) \xrightarrow{L} Y$.

Beweis.

(i) $\Rightarrow$ (ii): Ergibt sich durch Anwendung des Stetigkeitssatzes 1.3.3, wobei lediglich zu beachten ist, daß gemäß Meßbarkeitsvoraussetzung (M3) für jedes $(x_1, \dots, x_k) \in X^k, k \in \mathbb{N}$ die Abbildung $g: S \rightarrow \mathbb{R}^k, g(f) := (f(x_1), \dots, f(x_k))$, $A, \mathcal{B}_k^*$ -meßbar ist (die d-Stetigkeit ist klar).

(ii) $\Rightarrow$ (i): Aus der Konvergenz der endlichdimensionalen Randverteilungen folgt sofort, daß für zwei mögliche Häufungspunkte X_0^1, X_0^2 der Folge $(X_n)_{n \in \mathbb{N}}$ (d.h. es existieren Teilfolgen, die L_b -konvergent sind gegen X_0^1 bzw. X_0^2) gilt: $L(X_0^1)$ und $L(X_0^2)$ stimmen auf $S_0 \cap \sigma(\{\pi_x(S) : x \in X\})$ überein. Da zusätzlich jedoch die Verteilungen von X_0^1 bzw. X_0^2 auf S_0 konzentriert sind, folgt: $L(X_0^1)$ und $L(X_0^2)$ stimmen auch auf $\mathcal{B}_b(S)$ überein

das Beispiel auf Seite 111/112 in GÄNSSLER(1983). Wir werden nun hinreichende Bedingungen für die relative l_b -Folgenkompaktheit kennenlernen, die eine Mischung darstellen aus Forderungen an die "Zuwächse" der Zufallselemente $X_n, n \in \mathbb{N}$, und das Anwachsen der Funktion $N(\epsilon, X, s)$ (bei fallendem ϵ). Ein ähnliches Resultat im Fall von Partialsummen von unabhängigen stochastischen Prozessen findet man bei ALEXANDER(1984b), Theorem 2.2.

1.4.3 Theorem. Es gelten die Bedingungen (1) und (2), d.h.:

- (1) $\mathbb{P}(|X_n(x) - X_n(y)| \geq \epsilon) \leq f(s(x, y), n, \epsilon)$ für jedes $x, y \in X$, $\epsilon > 0$ sowie $n \in \mathbb{N}$, wobei $f: (0, \infty)^3 \rightarrow [0, \infty)$ eine in der ersten Komponente monoton wachsende, in der zweiten Komponente monoton fallende Funktion ist.
- (2) Es existieren Funktionen $r: (0, \infty) \rightarrow (0, \infty)$ sowie $g: (0, \infty) \rightarrow (0, \infty)$ mit:

- (i) r ist strikt fallend, stetig und $\lim_{t \rightarrow 0} r(t) = \infty$.
- (ii) g ist monoton wachsend mit $\int_0^2 t^{-1} \cdot g(t) dt < \infty$.
- (iii) $\int_0^1 t^{-1} \cdot N(\frac{t}{2}, X, s)^2 \cdot f(6t, r(4t), \epsilon g(t)) dt < \infty$ für $\epsilon > 0$.

Dann folgt für jedes $0 < \delta \leq 1$, $0 < \epsilon$ sowie $n \geq r(\delta)$:

$$\begin{aligned} \mathbb{P}^*(\omega_{X_n}(\delta) \geq 5\epsilon) &\leq 2 \cdot \int_{a_n}^{\delta} t^{-1} \cdot N(\frac{t}{2}, X, s)^2 \cdot f(6t, r(4t), \frac{\epsilon}{2K_1} \cdot g(t)) dt \\ &\quad + N(\delta, X, s)^2 \cdot f(5\delta, n, \epsilon) \\ &\quad + \mathbb{P}^*(\omega_{X_n}(b_n) \geq \epsilon). \end{aligned}$$

Dabei ist $K_1 \geq \int_0^2 t^{-1} \cdot g(t) dt$, $r(b_n) = n$ sowie $a_n = \frac{1}{4} b_n$.

Beweis. Sei $0 < \delta \leq 1$, $0 < \varepsilon$ beliebig. Sei weiter $n \geq r(\delta)$. Wir setzen für jedes $k \in \mathbb{N}$ $\delta_k := \delta \cdot 2^{1-k}$ und erhalten aus der Voraussetzung (V1): zu jedem $k \in \mathbb{N}$ existiert eine endliche Teilmenge E_k von X mit den Eigenschaften:

(α) Zu jedem $x \in X$ gibt es ein $y(x, k) \in E_k$ mit $s(x, y(x, k)) \leq 2\delta_k$.

(β) $|E_k|$ ist minimal unter allen Teilmengen von X mit

Eigenschaft (α), d.h. $|E_k| = N(\delta_k, X, s) =: N_k$.

Dann gilt:

$$\begin{aligned} \omega_{X_n}(\delta) &= \sup\{|X_n(x) - X_n(z)| : s(x, z) \leq \delta\} \\ &\leq \sup\{|X_n(x) - X_n(y(x, 1))| + |X_n(y(x, 1)) - X_n(y(z, 1))| + \\ &\quad |X_n(y(z, 1)) - X_n(z)| : s(x, z) \leq \delta\} \\ &\leq 2 \cdot \sup\{|X_n(x) - X_n(y(x, 1))| : x \in X\} + \\ &\quad \sup\{|X_n(x) - X_n(z)| : x, z \in E_1 \text{ und } s(x, z) \leq 5\delta\} \\ &=: 2 \cdot S_n^1 + S_n^2 \end{aligned}$$

Mit Hilfe eines "Chaining"-Arguments wollen wir zunächst den Term S_n^1 abschätzen. Mit $m(n) := \sup\{k \in \mathbb{N} : r(2\delta_k) \leq n\}$ (ist vernünftig definiert wegen $r(\delta_1) \leq n$ und $\lim_{t \rightarrow 0} r(t) = \infty$) folgt:

$$\begin{aligned} S_n^1 &\leq \sup\{\sum_{k=1}^{m(n)} |X_n(y(x, k)) - X_n(y(x, k+1))| + \\ &\quad |X_n(y(x, m(n)+1)) - X_n(x)| : x \in X\} \\ &\leq \sum_{k=1}^{m(n)} \sup\{|X_n(y(x, k)) - X_n(y(x, k+1))| : x \in X\} + \\ &\quad \omega_{X_n}(2\delta_{m(n)+1}) \\ &=: S_n^3 + \omega_{X_n}(2\delta_{m(n)+1}) \end{aligned}$$

Insgesamt erhalten wir somit (da aufgrund der Voraussetzung

$$(2)(i) \quad 2\delta_{m(n)+1} < b_n):$$

$$(+)\quad \omega_{X_n}(\delta) \leq 2 \cdot S_n^3 + S_n^2 + 2 \cdot \omega_{X_n}(b_n).$$

Für jedes $k \in \mathbb{N}$ sei nun $\tilde{\varepsilon}_k := g(\delta_k)$. Dann gilt wegen der Voraussetzung (2)(ii) (mit $\delta_0 := 2\delta$):

$$0 < \sum_{k \in \mathbb{N}} \tilde{\varepsilon}_k = 2 \cdot \sum_{k \in \mathbb{N}} \int_{\delta_k}^{\delta_{k-1}} \delta_{k-1}^{-1} g(\delta_k) dt \leq 2 \cdot \int_0^{2\delta} t^{-1} g(t) dt \leq 2K_1.$$

Mit $\varepsilon_k := \frac{\varepsilon}{2K_1} \cdot \tilde{\varepsilon}_k$, $k \in \mathbb{N}$, folgt nun:

$$\begin{aligned} \mathbb{P}(S_n^3 \geq \varepsilon) &\leq \mathbb{P}\left(S_n^3 \geq \sum_{k=1}^{m(n)} \varepsilon_k\right) \\ &\leq \sum_{k=1}^{m(n)} \mathbb{P}\left(\sup\{|X_n(x) - X_n(z)| : x, z \in E_k \cup E_{k+1}; s(x, z) \leq 3\delta_k\} \geq \varepsilon_k\right) \\ &\leq \sum_{k=1}^{m(n)} N_k \cdot N_{k+1} \cdot \sup\{\mathbb{P}(|X_n(x) - X_n(z)| \geq \varepsilon_k) : x, z \in X; s(x, z) \leq 3\delta_k\} \\ &\leq \sum_{k=1}^{m(n)} 2 \int_{\delta_{k+1}}^{\delta_k} \delta_k^{-1} N_{k+1}^2 f(3\delta_k, n, \frac{\varepsilon}{2K_1} g(t)) dt \\ &\leq \sum_{k=1}^{m(n)} 2 \int_{\delta_{k+1}}^{\delta_k} t^{-1} N(\frac{t}{2}, X, s)^2 f(6t, r(2\delta_k), \frac{\varepsilon}{2K_1} g(t)) dt \\ &\leq 2 \cdot \int_{a_n}^{\delta} t^{-1} N(\frac{t}{2}, X, s)^2 f(6t, r(4t), \frac{\varepsilon}{2K_1} g(t)) dt \end{aligned}$$

Da außerdem gilt:

$$\mathbb{P}(S_n^2 \geq \varepsilon) \leq N(\delta, X, s)^2 f(5\delta, n, \varepsilon), \text{ folgt wegen (+) unmittelbar die Behauptung.}$$

□

Ausgehend von Theorem 1.4.3 sind wir nun in der Lage, eine hinreichende Bedingung für die relative L_b -Folgenkompaktheit von $(X_n)_{n \in \mathbb{N}}$ herzuleiten.

1.4.4 Korollar. Es gelten die Bedingungen (1) und (2) aus

Theorem 1.4.3 sowie (mit $r(b_n) = \max\{n, r(1)\}$ für $n \in \mathbb{N}$):

$$(3) \limsup_{n \rightarrow \infty} \mathbb{P}^*(\omega_{X_n}(b_n) \geq \epsilon) = 0 \text{ für jedes } \epsilon > 0.$$

$$(4) \limsup_{n \rightarrow \infty} \mathbb{P}(|X_n(x)| \geq K) \rightarrow 0 \text{ für } K \uparrow \infty \text{ und ein } x \in X.$$

$$(5) \limsup_{n \rightarrow \infty} f(t, n, K) \rightarrow 0 \text{ für } K \uparrow \infty \text{ und alle } t \in [0, \infty).$$

Dann ist $(X_n)_{n \in \mathbb{N}}$ relativ L_b -folgenkompakt.

Beweis. Wir betrachten noch einmal den Beweis zu 1.4.3 und

übernehmen die dortigen Notationen. Für alle $0 < \epsilon, 0 < \delta \leq 1$ mit

$$g(\delta) \leq \epsilon \quad (\lim_{t \downarrow 0} g(t) = 0) \quad \text{sowie } n \geq r(\delta) \text{ gilt } (N(\cdot) := N(\cdot, X, s)):$$

$$\mathbb{P}(S_n^2 \geq \epsilon) \leq N(\delta)^2 \cdot f(5\delta, n, g(\delta)),$$

$$\text{also } \limsup_{n \rightarrow \infty} \mathbb{P}(S_n^2 \geq \epsilon) \leq N(\delta)^2 \cdot f(5\delta, r(4\delta), g(\delta)) =: G(\delta).$$

Da gemäß Bedingung (2)(iii) die Funktion $\delta^{-1} \cdot G(\delta)$ integrier-

bar ist, existiert eine Nullfolge $(\delta_m)_{m \in \mathbb{N}}$ mit $G(\delta_m) \rightarrow 0$

für $m \rightarrow \infty$. Insgesamt folgt wegen Theorem 1.4.3 und Beding-

ung (3):

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}^*(\omega_{X_n}(\delta) \geq 5\epsilon) \leq \lim_{m \rightarrow \infty} \sup G(\delta_m) = 0 \text{ für } \epsilon > 0.$$

Damit wäre also die Bedingung (iii)(A) aus Theorem 1.3.4

erfüllt. Der Nachweis von (iii)(B) hingegen bereitet keine

Probleme, denn es gilt für jedes $y \in X$ gemäß Voraussetzung (4), (5):

$$\begin{aligned} \limsup_{n \rightarrow \infty} \mathbb{P}(|X_n(y)| \geq K) &\leq \limsup_{n \rightarrow \infty} [\mathbb{P}(|X_n(y) - X_n(x)| \geq \frac{K}{2}) + \\ &\quad \mathbb{P}(|X_n(x)| \geq \frac{K}{2})] \\ &\leq \limsup_{n \rightarrow \infty} [f(s(x, y), n, \frac{K}{2}) + \mathbb{P}(|X_n(x)| \geq \frac{K}{2})] \\ &\rightarrow 0 \text{ für } K \uparrow \infty \text{ und alle } y \in X. \end{aligned}$$

□

1.4.5 Bemerkung.

- (i) Die Bedingung (2)(iii) in Korollar 1.4.4 (bzw. in Theorem 1.4.3) ist im Hinblick auf empirische C -Prozesse zugeschnitten auf ein im wesentlichen polynomiales Verhalten von $N(.,X,s)$ (was z.B. im Fall " C ist Vapnik-Chervonenkis-Klasse" gesichert ist; vgl. hierzu (VC3) in Kapitel II) in Verbindung mit einer geeigneten Exponentialungleichung (z.B. vom Bernsteintyp). Es wird also bei Anwendung von 1.4.4 in erster Linie darauf ankommen, die Bedingung (3) sicherzustellen, d.h. das Verhalten von X_n auf "sehr kleinen Mengen" (von der Größenordnung b_n ; $n=r(b_n)$) zu untersuchen - vgl. hierzu auch GINÉ-ZINN(1984), (1.2). Überdies ist die Bedingung (4) bei empirischen C -Prozessen trivialerweise immer erfüllt durch die Forderung " $\emptyset \in C$ ".
- (ii) Theorem 1.4.3 bzw. Korollar 1.4.4 entsprechen in gewisser Weise Corollary 1.19 in GÄNSSLER-HÄUSLER-SCHNEEMEIER(1984).
- (iii) Das "Chaining"-Argument im Beweis zu Theorem 1.4.3 findet man in entsprechender Form bereits bei DUDLEY(1978); vgl. auch ALEXANDER(1984a), GINÉ-ZINN(1984) sowie GÄNSSLER(1985).
- (iv) Korollar 1.4.4 eignet sich nicht nur zur Beschreibung empirischer C - und F -Prozesse, sondern allgemein für alle stochastischen Prozesse, die die Entropiebedingung an den Parameterraum bzw. die Forderung an die Zuwächse aus 1.4.3 erfüllen.
- (v) Im weiteren Verlauf setzen wir stets o.E. voraus, daß auch die Grenzelemente X_0 über $(\Omega, \mathcal{F}, \mathbb{P})$ definiert seien und außerdem gilt: $X_0(x)$ ist $\mathcal{F}, \mathcal{B}^*$ -meßbar für alle $x \in X$.

KAPITEL II: EIN FUNKTIONALER GRENZWERTSATZ FÜR DEN "GEBOOT-STRAPPTEN" EMPIRISCHEN C-PROZESS INDIZIERT NACH VAPNIK-CHERVONENKIS-KLASSEN VON MENGEN

Im Verlauf des Kapitels II wollen wir stets die folgende Situation betrachten:

Sei $(\xi_{ni})_{i \leq i(n), n \in \mathbb{N}}$ ein Dreiecksschema von zeilenweise unabhängigen (nicht notwendig identisch verteilten) Zufallselementen in einen meßbaren Raum $(Y, \mathcal{B})$, d.h. es existiert ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ so, daß für jedes $n \in \mathbb{N}$ und $i \leq i(n)$ gilt: $\xi_{ni}: \Omega \rightarrow Y$ ist $\mathcal{F}, \mathcal{B}$ -meßbar. Sei weiter $C \subset \mathcal{B}$ eine abzählbare Vapnik-Chervonenkis-Klasse mit $\emptyset \in C$, d.h. es existiert ein $V(C) \in \mathbb{N}$ mit:

- (i) $|C \cap \{y_1, \dots, y_{V(C)}\}| < 2^{V(C)}$ für alle $y_1, \dots, y_{V(C)} \in Y$.
- (ii) $V(C)$ ist minimal gewählt mit Eigenschaft (i).

Dann definieren wir für jedes $C \in \mathcal{B}$ und $n \in \mathbb{N}$:

$$\mu_n(C) := i(n)^{-1} \cdot \sum_{i=1}^{i(n)} 1_C(\xi_{ni}),$$

$$\nu_n(C) := i(n)^{-1} \cdot \sum_{i=1}^{i(n)} \mathbb{P}(\xi_{ni} \in C),$$

$$\beta_n(C) := i(n)^{\frac{1}{2}} \cdot [\mu_n(C) - \nu_n(C)].$$

Den stochastischen Prozeß $\beta_n = (\beta_n(C))_{C \in \mathcal{C}}$ bezeichnet man für jedes $n \in \mathbb{N}$ als empirischen C-Prozeß, und unser Ziel wird es sein, ihn auf Verteilungskonvergenz hin zu untersuchen. Diese werden wir natürlich im Rahmen der L_b -Theorie formulieren, wobei der Raum X aus Kapitel I das Mengensystem $\mathcal{C}$ sein wird, das wir mit einer geeigneten Halbmetrik s versehen. Ausgehend

von den Ergebnissen aus Kapitel I wollen wir dann einen funktionalen Grenzwertsatz für den sog. gebootstrappten empirischen C -Prozeß herleiten.

Zunächst wollen wir jedoch einige grundlegende Eigenschaften von (nicht notwendig abzählbaren) Vapnik-Chervonenkis-Klassen bereitstellen. Dazu betrachten wir für jedes $n \in \mathbb{N}$ die Halbmetrik

$$s_n(C, D) := v_n(C \Delta D) \text{ für } C, D \in \mathcal{C}$$

auf $\mathcal{C}$ und definieren $N(\cdot, \mathcal{C}, s_n)$ wie in 1.4.2. Dann gilt:

(VC1) Für beliebige $y_1, \dots, y_k \in Y$ gilt:

$|\mathcal{C} \cap \{y_1, \dots, y_k\}| \leq k^{V(\mathcal{C})} + 1$, wobei $k \in \mathbb{N}$ beliebig gewählt ist (vgl. zu dieser sogenannten Vapnik-Chervonenkis-Ungleichung GÄNSSLER(1983), S. 27, Lemma 9).

(VC2) Mit $\mathcal{C}$ ist auch $\mathcal{C}^- := \{C \setminus D : C, D \in \mathcal{C}\}$ eine Vapnik-Chervonenkis-Klasse.

(VC3) Für jedes $n \in \mathbb{N}$ gilt: $N(x, \mathcal{C}, s_n) \leq K \cdot x^{-V(\mathcal{C})} \cdot |\log x|^{V(\mathcal{C})}$ für alle $0 < x \leq \frac{1}{2}$, wobei K lediglich von $V(\mathcal{C})$ abhängt (dies folgt mit Lemma (7.13), DUDLEY(1978)).

Bemerkung:

(i) Die Eigenschaft (VC1) werden wir benutzen, um in Korollar 1.4.4 die Bedingung (3) zu verifizieren. Die Gültigkeit von (VC3) hingegen wird uns in die Lage versetzen, die Forderung (2)(iii) aus Theorem 1.4.3 sicherzustellen, wobei wir für die dortige Bedingung (1) (Abschätzung der "Zuwächse" von β_n) von der Bernsteinungleichung Gebrauch machen.

(ii) Die Forderung der Abzählbarkeit von $\mathcal{C}$ findet man auch bei POLLARD(1982). Dies stellt jedoch keine wesentliche Ein-

schränkung dar, da sehr viele Prozesse von Interesse auch "separabel" sind (in dem Sinne, daß sie bereits eindeutig bestimmt sind durch ihre Werte auf einer abzählbaren Indexmenge). Es sei an dieser Stelle auch ausdrücklich darauf hingewiesen, daß diese Voraussetzung lediglich dazu dient, die Meßbarkeit der in Kapitel II auftretenden Größen zu sichern. Die Abzählbarkeit von C hätten wir also auch durch eine geeignete Meßbarkeitsforderung ersetzen können, ohne daß sich dadurch die Beweise ändern würden.

2.1 Stochastische Ungleichungen

In diesem Abschnitt wollen wir stochastische Ungleichungen herleiten, die uns bei festem $n \in \mathbb{N}$ die Kontrolle von β_n gewährleisten. Für den gesamten Verlauf von 2.1 wollen wir davon ausgehen, daß für jedes $n \in \mathbb{N}$ gilt: $i(n) = n$. Dann gilt zunächst:

2.1.1 Lemma. Sei $R \subset \mathcal{B}$ abzählbar, $n \in \mathbb{N}$ und $k \in \{1, \dots, n\}$ fest vorgegeben. Dann gilt für alle $z \in \mathbb{N}$ mit $z \geq 2$, falls

$$\alpha := \sup\{\nu_n(C) : C \in R\} < \frac{1}{n} :$$

$$\mathbb{P}(\sup\{\mu_n(C) : C \in R\} \geq \frac{k}{n}) \leq (1-\alpha)^{1-z} \sup\{\sum_{i=k}^n m_i(R, \bar{y}) \cdot z^{-i} : \bar{y} \in Y^{zn}\},$$

wobei für jedes $i \in \{0, \dots, n\}$ und $\bar{y} = (y_1, \dots, y_{zn}) \in Y^{zn}$ definiert

$$\text{sei: } m_i(R, \bar{y}) := |\{C \cap \{y_1, \dots, y_{zn}\} : C \in R \text{ und } \sum_{1 \leq j \leq zn} 1_C(y_j) = i\}|.$$

Bemerkung. Die Definition der "Wachstumsfunktionen" $m_i(R, \cdot)$ ist natürlich eng an die Voraussetzung " R ist Vapnik-Chervonenkis-Klasse" geknüpft, da ja für jedes $\bar{y} \in Y^{zn}$ gemäß (VC1)

dann gilt: $\sum_{i=0}^n m_i(R, \bar{y}) = |R \cap \{y_1, \dots, y_{zn}\}| \leq (zn)^{V(R)} + 1.$

Falls hingegen R keine Vapnik-Chervonenkis-Klasse ist, steht auf der rechten Seite von 2.1.1 im allgemeinen keine vernünftige Abschätzung.

Beweis zu Lemma 2.1.1. Sei $R \subset \mathcal{B}$ abzählbar, $n \in \mathbb{N}$ und $k \in \{1, \dots, n\}$ fest vorgegeben. Sei weiter $z \in \mathbb{N}$ mit $z \geq 2$. Dann existieren unabhängige Zufallselemente $(\xi_1^j, \dots, \xi_n^j)$, $j = 1, \dots, z$ mit $L((\xi_1^j, \dots, \xi_n^j)) = L((\xi_{n1}, \dots, \xi_{nn}))$ für alle $j = 1, \dots, z$ (wobei o.E. jedes ξ_i^j über dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ definiert und $\mathcal{F}, \mathcal{B}$ -meßbar sei).

Für jedes $j = 1, \dots, z$ und $\bar{y} = (y_1, \dots, y_n) \in Y^n$ definieren wir:

$$\mu_n^j(C) := n^{-1} \cdot \sum_{i=1}^n 1_C(\xi_i^j) \quad \text{für } C \in \mathcal{B},$$

$$S(\bar{y}) := \{C \in \mathcal{R} : \sum_{i=1}^n 1_C(y_i) \geq k\}.$$

Falls dann $S(\bar{y}) \neq \emptyset$, setzen wir $C(\bar{y}) \in S(\bar{y})$ beliebig.

Falls hingegen $S(\bar{y}) = \emptyset$, setzen wir $C(\bar{y}) := \emptyset$.

Damit folgt, wobei $\bar{\xi} = (\xi_1^1, \dots, \xi_n^1)$:

$$\mathbb{P}(\bigcup_{C \in \mathcal{R}} \{ \mu_n^1(C) \geq \frac{k}{n} \text{ und } \mu_n^j(C) = 0 \text{ für alle } j = 2, \dots, z \}) \stackrel{\text{Fubini}}{=}$$

$$\int 1(S(\bar{y}) \neq \emptyset) \cdot \mathbb{P}(\bigcup_{C \in S(\bar{y})} \{ \mu_n^j(C) = 0 \text{ für alle } j = 2, \dots, z \}) Q_{\bar{\xi}}(d\bar{y}) \leq$$

$$b_n \cdot \int 1(S(\bar{y}) \neq \emptyset) Q_{\bar{\xi}}(d\bar{y}) = b_n \cdot \mathbb{P}(\bigcup_{C \in \mathcal{R}} \{ \mu_n(C) \geq \frac{k}{n} \})$$

$$\text{mit } b_n := \inf \{ \mathbb{P}(\mu_n^j(C(\bar{y})) = 0 \text{ für alle } j = 2, \dots, z) : \bar{y} \in Y^n \}.$$

Mit Hilfe der Markoff-Ungleichung erhalten wir:

$$b_n \geq \inf \left\{ \prod_{j=2}^z (1 - \mathbb{P}(\mu_n^j(C) \geq \frac{1}{n})) : C \in \mathcal{R} \right\}$$

$$\geq \inf \left\{ \prod_{j=2}^z (1 - n \mathbb{E}(\mu_n^j(C))) : C \in \mathcal{R} \right\}$$

$$= \inf \left\{ (1 - n \nu_n(C))^{z-1} : C \in \mathcal{R} \right\} \geq (1 - na)^{z-1},$$

d.h. es folgt insgesamt:

$$\mathbb{P}(\sup\{\mu_n(C) : C \in \mathcal{R}\} \geq \frac{k}{n}) \leq (1-na)^{1-z} \cdot \sum_{s=k}^n a_s, \text{ wobei f\"ur jedes}$$

$$s = k, \dots, n \quad a_s := \mathbb{P}(\bigcup_{C \in \mathcal{R}} \{\mu_n^1(C) = \frac{s}{n} \text{ und } \mu_n^j(C) = 0 \text{ f\"ur alle } j=2, \dots, z\}).$$

Wenn wir nun noch f\"ur jedes $s = k, \dots, n$ a_s geeignet absch\"atzen

k\"onnen, w\"are der Beweis zu Lemma 2.1.1 beendet. Sei also im

folgenden s fest gew\"ahlt. Dann definieren wir:

$$\mathcal{R}(s, \bar{y}) := \{C \cap \{y_1^1, \dots, y_n^1, \dots, y_1^z, \dots, y_n^z\} : C \in \mathcal{R}, \sum_{i=1}^n \sum_{j=1}^z 1_C(y_i^j) = s\}$$

$$\text{f\"ur } \bar{y} = (y_1^1, \dots, y_n^1, \dots, y_1^z, \dots, y_n^z) \in Y^{zn}.$$

Mit $\bar{\xi} = (\xi_1^1, \dots, \xi_n^1, \dots, \xi_1^z, \dots, \xi_n^z)$ gilt zun\"achst:

$$\begin{aligned} a_s &\leq \int \sum_{C \in \mathcal{R}(s, \bar{y})} 1\left(\sum_{i=1}^n 1_C(y_i^1) = s \text{ und } \sum_{i=1}^n 1_C(y_i^j) = 0 \text{ f\"ur } j=2, \dots, z\right) Q_{\bar{\xi}}(d\bar{y}) \\ &= \int \sum_{C \in \mathcal{R}(s, \bar{y})} 1\left(\sum_{i=1}^n 1_C(\pi_i(y_i^1)) = s, \sum_{i=1}^n 1_C(\pi_i(y_i^j)) = 0 \text{ f\"ur } j=2, \dots, z\right) Q_{\bar{\xi}}(d\bar{y}) \end{aligned}$$

f\"ur jede Wahl von Permutationen $\pi_1, \dots, \pi_n \in S(z) := \{\text{Permu-}$

tationen von $\{1, \dots, z\}\}$, wobei f\"ur alle $i=1, \dots, n$ sowie

$j=1, \dots, z$ $\pi_i(y_i^j) := y_i^r$ mit $r = \pi_i(j)$ (f\"ur beliebiges $\bar{y} \in Y^{zn}$).

Hierbei wurde lediglich benutzt, da\ss f\"ur alle $\pi_1, \dots, \pi_n \in S(z)$

gilt: $L(\bar{\xi}) = L((\pi_1(\xi_1^1), \dots, \pi_n(\xi_n^1), \dots, \pi_1(\xi_1^z), \dots, \pi_n(\xi_n^z)))$.

Hieraus ergibt sich nun wegen $|S(z)| = z!$:

$$a_s \leq \int \sum_{C \in \mathcal{R}(s, \bar{y})} f(C, \bar{y}) Q_{\bar{\xi}}(d\bar{y}), \text{ wobei f\"ur } C \in \mathcal{R}(s, \bar{y}) \text{ und } \bar{y} \in Y^{zn}$$

$$f(C, \bar{y}) := (z!)^{-n} \cdot \sum_{\pi_1, \dots, \pi_n \in S(z)} 1 \left(\sum_{i=1}^n 1_{C(\pi_i(y_i^1))} = s, \sum_{i=1}^n 1_{C(\pi_i(y_i^j))} = 0 \right.$$

für alle $j = 2, \dots, z$).

Für festes $\bar{y} \in Y^{zn}$ und $C \in R(s, \bar{y})$ gilt aber: Falls $f(C, \bar{y}) > 0$,

existiert $(\hat{\pi}_1, \dots, \hat{\pi}_n) \in S(z)^n$ mit $\sum_{i=1}^n 1_{C(\hat{\pi}_i(y_i^1))} = s$

sowie $\sum_{i=1}^n 1_{C(\hat{\pi}_i(y_i^j))} = 0$ für $j = 2, \dots, z$. Sei also o.E.

$$1_{C(\hat{\pi}_1(y_1^1))} = 1_{C(\hat{\pi}_2(y_2^1))} = \dots = 1_{C(\hat{\pi}_s(y_s^1))} = 1 \text{ und}$$

$$1_{C(\hat{\pi}_{s+1}(y_{s+1}^1))} = 1_{C(\hat{\pi}_n(y_n^1))} = 0 = 1_{C(\hat{\pi}_i(y_i^j))} \text{ für alle}$$

$i = 1, \dots, n$ sowie $j = 2, \dots, z$. Es ergibt sich unmittelbar:

$$\begin{aligned} f(C, \bar{y}) &= (z!)^{-n} \cdot |\{\pi_1, \dots, \pi_n \in S(z)^n : \pi_i(1) = \hat{\pi}_i(1) \\ &\quad \text{für } i = 1, \dots, s\}| \\ &= (z!)^{-n} [(z-1)!]^s (z!)^{n-s} \\ &= z^{-s}. \end{aligned}$$

Damit folgt die Behauptung. $\square$

2.1.2 Korollar. Für alle $n \in \mathbb{N}$ mit $n \geq V(C^-) + 1$, $x > 0$ und

$k \in \{V(C^-) + 1, \dots, n\}$ gilt:

$$\mathbb{P}(\sup\{\mu_n(C) : C \in C^- \text{ mit } v_n(C) \leq \frac{1}{2} \cdot n^{-\frac{k+x}{k-V(C^-)}}\} \geq \frac{k}{n}) \leq 4 \cdot n^{-x}.$$

Beweis. Dies ist eine unmittelbare Folgerung aus Lemma 2.1.1,

denn mit $v := V(C^-)$, $s := \frac{v+x}{k-v}$ sowie $z := \lfloor n^s \rfloor + 1$ erhalten wir (wo-

bei $\lfloor \cdot \rfloor$ jetzt wie auch im weiteren Verlauf stets die Gaußklammer bezeichne):

$$\mathbb{P}(\sup\{\mu_n(C) : C \in C^- \text{ mit } v_n(C) \leq \frac{1}{2} \cdot n^{-1-s}\} \geq \frac{k}{n}) \stackrel{(VC1)}{\leq}$$

$$\begin{aligned}
 &\leq 2 \cdot \left(1 - \frac{1}{2} n^{-s}\right)^{1-z} \cdot (zn)^v \cdot z^{-k} \\
 &\leq 2 \cdot \left(1 - \frac{1}{2} n^{-s} \cdot (z-1)\right)^{-1} \cdot n^v \cdot z^{v-k} \\
 &\leq 2 \cdot \left(1 - \frac{1}{2}\right)^{-1} \cdot n^v \cdot n^{-v-x} = 4 \cdot n^{-x}. \quad \square
 \end{aligned}$$

2.1.3 Bemerkung.

(i) Im Beweis zu 2.1.1 konnten im Gegensatz zum Beweis der Vapnik-Chervonenkis-Ungleichung (vgl. Lemma 10 in GÄNSSLER(1983)) nicht alle Permutationen von $(1, \dots, zn)$ benutzt werden, da wir keine "i.i.d."-Voraussetzung gemacht haben. Weiterhin ist bemerkenswert, daß der Beweis ohne Abschätzungen für die "Tails" der Hypergeometrischen Verteilung oder der Binomialverteilung auskommt (im Gegensatz zum Beweis von Lemma 10 in GÄNSSLER(1983) oder Korollar 2.9 in ALEXANDER(1984a)). Dies liegt natürlich daran, daß wir für die $z-1$ unabhängigen "Kopien" von $(\xi_{n1}, \dots, \xi_{nn})$ stets das Ereignis $\{1_C(\xi_i^j) = 0 \text{ für } i=1, \dots, n; j=2, \dots, z\}$ hinzugezogen haben, was uns eine exakte Berechnung der auftretenden Funktion $f(.,.)$ ermöglichte.

(ii) Lemma 2.1.1 ist selbstverständlich eine Aussage für "kleine" Mengen. Man kann sie z.B. vergleichen mit folgender Aussage: "Für unabhängige und über $(0,1)$ gleichverteilte Zufallsvariable $U_1, \dots, U_n$ gilt: $\mathbb{P}(U_i \geq \frac{x}{n} \text{ für alle } i=1, \dots, n) = (1 - \frac{x}{n})^n \rightarrow \exp(-x)$ für $n \rightarrow \infty$ und alle $x \in (0,1)$. Mit gegen 1 gehender Wahrscheinlichkeit werden wir also im Intervall $(0, n^{-2})$ keine Beobachtungen machen."

Daß man dieses "keine" für beliebige Vapnik-Chervonenkis-Klassen nicht mehr erwarten darf, mag das folgende Beispiel illustrieren:

Sei $C_R = \{C \subset (0,1) : |C| \leq R\}$ für $R \in \mathbb{N}$, $U_1, \dots, U_n$ wieder wie eben gewählt mit $n \geq R$. Dann gilt für jedes $\varepsilon > 0$ (wobei λ das Lebesgue-Maß bezeichne und μ_n die empirische Verteilung zu $U_1, \dots, U_n$):

(*) $\sup\{\mu_n(C) : \lambda(C) \leq \varepsilon, C \in C_R\} = \frac{R}{n}$, C_R ist eine Vapnik-Chervonenkis-Klasse mit $C_R = C_R^-$ und $V(C_R) = R + 1$.

Dieses Beispiel zeigt auch, daß Korollar 2.1.2 im Vergleich zu (*) höchstens um "eine Beobachtung" verbessert werden könnte, da insbesondere folgt (in der Situation von 2.1.2):

$\mathbb{P}\left(\sup\{\mu_n(C) : C \in C^-, \mu_n(C) \leq \frac{1}{2} \cdot n^{-V(C^-)-2}\} \geq \frac{V(C^-)+1}{n}\right) \leq 4n^{-1}$
für alle $n \in \mathbb{N}$ mit $n \geq V(C^-)+1$.

(iii) 2.1.1 kann auch im Zusammenhang mit ALEXANDER(1984a) gesehen werden. Im dortigen Korollar 2.9 wird zwar keine Bedingung an α (α wie in 2.1.1 definiert) gestellt, dafür muß man jedoch fordern: $k^2 \geq M \cdot \log n$ für eine nur von C abhängige Konstante M . Lemma 2.1.1 liefert also unter stärkerer Bedingung an α eine Abschwächung dieses Resultats.

Das folgende Theorem liefert nun eine Abschätzung für den Stetigkeitsmodul von β_n bei festem $n \in \mathbb{N}$.

2.1.4 Theorem. Für alle $0 < s < \frac{1}{2}$ existieren positive Konstanten K und R , die lediglich von s und C abhängen, mit der folgenden Eigenschaft:

(i) Für alle $0 < \delta \leq 1$, $0 < \varepsilon$ und $n \geq \max\{(2\delta)^{-1}, 4\varepsilon^{-2}\}$ gilt:

$$\mathbb{P}(\tilde{\omega}_{\beta_n}(\delta) \geq 5\varepsilon) \leq K[\delta^{-1-2V(C)} \gamma^{-d} + \varepsilon^{-2V(C^-)}] \exp(-R\gamma),$$

wobei $\gamma := \eta^2(1+\eta)^{-1}$ mit $\eta := \varepsilon \delta^{\frac{s-1}{2}}$, $d := \frac{2+4V(C)}{1-2s}$ und

$$\tilde{\omega}_{\beta_n}(\delta) := \sup\{|\beta_n(C) - \beta_n(D)| : C, D \in C \text{ mit } \nu_n(C \Delta D) \leq \delta\}.$$

(ii) Für alle $0 < \delta \leq 1$ und $n \geq (2\delta)^{-1}$ gilt:

$$\delta^{1+2V(C^-)} \mathbb{E}(\exp(R\delta^{s-\frac{1}{2}} \tilde{\omega}_{\beta_n}(\delta))) \leq K.$$

Beweis. Sei $0 < s \leq \frac{1}{2}$ fest vorgegeben. Im Verlauf des gesamten Beweises zu (i) bezeichnen K und R beliebige Konstanten, deren jeweiliger Wert variabel ist, jedoch immer nur von s und C abhängt (der exakte Wert spielt für unsere Zwecke keine Rolle).

(i) Sei $0 < \delta \leq 1$, $0 < \varepsilon$ und $n \geq \max\{(2\delta)^{-1}, 4\varepsilon^{-2}\}$ fest gewählt.

Dann betrachten wir die Halbmetrik

$$s(C, D) := v_n(C \Delta D) \text{ für } C, D \in C$$

auf C . In der Terminologie von Kapitel I ist dann $\omega_{\beta_n} = \tilde{\omega}_{\beta_n}$, und wir werden den Beweis natürlich mit Hilfe von Theorem 1.4.3 führen. Zunächst einmal gilt gemäß Bernsteinungleichung (vgl. GÄNSSLER(1983), S.9) für alle $z > 0$ und $C, D \in C$:

$$\mathbb{P}(|\beta_n(C) - \beta_n(D)| \geq z) \leq 2 \exp\left(-\frac{z^2}{2s(C, D) + \frac{4}{3}zn^{-1/2}}\right).$$

Mit $f(x, y, z) := 2 \exp\left(-\frac{z^2}{2x + \frac{4}{3}zy^{-1/2}}\right)$ für $x \geq 0$, $y > 0$ sowie $z > 0$

(für $y=0$ bzw. $z=0$ ist die Definition von f beliebig) ist dann die Bedingung (1) aus Theorem 1.4.3 erfüllt. Weiter betrachten wir Funktionen g und r wie folgt:

$$g(t) := t^s \text{ sowie } r(t) := (2t)^{-1} \text{ für } t > 0.$$

Dann ist unmittelbar klar, daß mit diesen Definitionen die Bedingungen (2)(i) bzw. (2)(ii) aus 1.4.3 sichergestellt sind mit

$$\int_0^2 t^{-1} \cdot g(t) dt = s^{-1} \cdot 2^s =: K_1.$$

Darüberhinaus folgt für alle $z > 0$ (wegen (VC3)):

$$\begin{aligned} & \int_0^1 t^{-1} N\left(\frac{t}{2}, C, s\right)^2 f(6t, r(4t), z g(t)) dt \\ & \leq \int_0^1 K t^{-2-2V(C)} \exp\left(-\frac{z^2 t^{2s}}{12t + \frac{4}{3} z t^s (8t)^{1/2}}\right) dt \\ & \leq \int_0^1 K t^{-2-2V(C)} \exp\left(-\frac{z^2}{12+4z} t^{s-\frac{1}{2}}\right) dt < \infty. \end{aligned}$$

Die Voraussetzungen von 1.4.3 sind also allesamt erfüllt und es folgt wegen $n \geq (2\delta)^{-1} = r(\delta)$:

$$\begin{aligned} \mathbb{P}(\tilde{\omega}_{\beta_n}(\delta) \geq 5\epsilon) & \leq K \cdot \left[\int_0^\delta t^{-2-2V(C)} \exp\left(-\frac{\epsilon^2 (2K_1)^{-2} t^{2s}}{12t + \frac{4}{3} [\epsilon (2K_1)^{-1} t^s] (8t)^{1/2}}\right) dt \right. \\ & \quad \left. + \delta^{-1-2V(C)} \exp\left(-\frac{\epsilon^2}{10\delta + \frac{4}{3} \epsilon n^{-1/2}}\right) \right] \\ & \quad + \mathbb{P}(\tilde{\omega}_{\beta_n}\left(\frac{1}{2n}\right) \geq \epsilon) =: K(S_1 + S_2) + S_3. \end{aligned}$$

Für S_2 gilt dann zunächst (wegen $n^{-1/2} \leq (2\delta)^{1/2}$):

$$\begin{aligned} S_2 & \leq \int_0^{2\delta} \delta^{-1} \delta^{-1-2V(C)} \exp\left(-\frac{\epsilon^2}{10t + 2\epsilon t^{1/2}}\right) dt \\ & \leq 2^{2+2V(C)} \cdot \int_0^{2\delta} t^{-2-2V(C)} \exp\left(-\frac{\epsilon^2 t^{s-1/2}}{10(2\delta)^{s+1/2} + 2\epsilon (2\delta)^s}\right) dt \end{aligned}$$

Hieraus folgt sofort (wobei γ und d wie in der Formulierung von 2.1.4 (i)):

$$\begin{aligned} K(S_1 + S_2) & \leq K \int_0^{2\delta} t^{-2-2V(C)} \exp\left(-\frac{\gamma}{48K_1^2} \left(\frac{t}{\delta}\right)^{s-1/2}\right) dt \\ & \leq \left[\left(\frac{t}{\delta}\right)^{-1-2V(C)} = u \right] \\ & \quad K \delta^{-1-2V(C)} \int_{2^{-1-2V(C)}}^{\infty} \exp\left(-\frac{\gamma}{48K_1^2} u^{1/d}\right) du \\ & \leq K \delta^{-1-2V(C)} \exp\left(-\frac{\gamma}{50K_1^2} 2^{s-1/2}\right) \gamma^{-d}. \end{aligned}$$

Falls wir nun noch eine entsprechende Abschätzung für S_3 herleiten können, wäre (i) bewiesen. Hierfür wird Lemma 2.1.1 von entscheidender Bedeutung sein, denn es gilt wegen $n \geq 4\varepsilon^{-2}$ ([] bezeichne wieder die Gaußklammer):

$$\mathbb{P}(\tilde{\omega}_{\beta_n}(\frac{1}{2n}) \geq \varepsilon) \leq 2 \lim_{u \rightarrow 0} \mathbb{P}(|\beta_n(C)| \geq \frac{\varepsilon}{2} - u \text{ für ein } C \in C^- \text{ mit } v_n(C) \leq \frac{1}{2n})$$

$$\leq 2 \mathbb{P}(\mu_n(C) \geq \frac{\varepsilon}{2} n^{-1/2} - \frac{1}{2n} \text{ für } C \in C^- \text{ mit } v_n(C) \leq \frac{1}{2n})$$

$$\leq 2 \mathbb{P}(\mu_n(C) \geq \frac{[\frac{\varepsilon}{4} n^{1/2}]}{n} \text{ für ein } C \in C^- \text{ mit } v_n(C) \leq \frac{1}{2n})$$

$$\stackrel{2.1.1}{\leq} 2(1 - \frac{1}{2})^{1-4} [(4n)^{V(C^-)} + 1] \exp(-[\frac{\varepsilon}{4} n^{1/2}])$$

mit $z=4$

$$\leq K n^{V(C^-)} \exp(-\frac{\varepsilon}{4} n^{1/2})$$

$$\leq K n^{V(C^-)} (\frac{\varepsilon}{8} n^{1/2})^{-2V(C^-)} \exp(-\frac{\varepsilon}{8} n^{1/2})$$

$$\leq K \varepsilon^{-2V(C^-)} \exp(-\frac{\gamma}{12}) \quad (\text{wegen } n \geq (2\delta)^{-1}).$$

Damit ist Teil (i) gezeigt.

(ii) Sei $0 < \delta \leq 1$ und $n \geq (2\delta)^{-1}$ fest vorgegeben. Gemäß Teil (i) existieren dann positive Konstanten K und R mit der Eigenschaft (wobei γ und d zu δ, ε und s gewählt sei wie in der Formulierung von Theorem 2.1.4 (i)):

$$(*) \quad \mathbb{P}(\tilde{\omega}_{\beta_n}(\delta) \geq 5\varepsilon) \leq K [\delta^{-1-2V(C)} \gamma^{-d} + \varepsilon^{-2V(C^-)}] \exp(-R\gamma) \text{ für } \varepsilon > 0 \text{ mit } 4\varepsilon^{-2} \leq n.$$

Für $\varepsilon \geq \exp(R\delta^{s-1/2} n^{-1/2})$ gilt $4(\frac{2 \log \varepsilon}{R} \delta^{1/2} - s)^{-2} \leq n$ und es folgt daher:

$$\mathbb{E}(\exp(\frac{R}{10} \delta^{s-1/2} \tilde{\omega}_{\beta_n}(\delta))) = 1 + \int_1^\infty \mathbb{P}(\tilde{\omega}_{\beta_n}(\delta) \geq 5(\frac{2 \log \varepsilon}{R} \delta^{1/2} - s)) d\varepsilon$$

$$\begin{aligned} (*) \quad & \leq \exp(R\delta^{s-1/2}n^{-1/2}) + \int_1^{\infty} \min\{1, 2K\delta^{-1-2V(C^-)}h(R,\epsilon)\} d\epsilon \\ & \leq \exp(2^{1/2}R\delta^s) + \delta^{-1-2V(C^-)} \int_1^{\infty} \min\{1, 2K h(R,\epsilon)\} d\epsilon. \end{aligned}$$

Dabei ist $h(R,\epsilon) := [1 + \frac{(2\log\epsilon)^2}{R^2 + 2R\log\epsilon}]^{-d} \exp(-\frac{(2\log\epsilon)^2}{R + 2\log\epsilon})$ und es ist

klar, daß $\int_1^{\infty} \min\{1, 2K h(R,\epsilon)\} d\epsilon$ eine positive und endliche

Konstante ist, deren Wert nur von R und K (also nur von s und C) abhängt. Dies beendet den Beweis von 2.1.4. $\square$

2.1.5 Bemerkung.

(i) Theorem 2.1.4 (i) hätten wir auch aus Korollar 2.9 in ALEXANDER(1984a) herleiten können. Dort wird mit wesentlich aufwendigeren Mitteln ein (allerdings) allgemeineres Resultat bewiesen, so daß im Hinblick auf funktionale Grenzwertsätze für empirische C -Prozesse der vorliegende Beweis effizienter ist (man beachte, daß insbesondere der Beweis zu 2.1.1 relativ elementar war; vgl. hierzu 2.1.3 (i)).

(ii) Von 2.1.4 (ii) werden wir erst in Kapitel III Gebrauch machen. Dabei werden wir ganz entscheidend ausnutzen, daß $\exp(Rx)$ für alle $R > 0$ eine monoton wachsende und konvexe Funktion darstellt.

Um nun auch bestmögliche stochastische Ungleichungen zur Herleitung von Glivenko-Cantelli-Aussagen zu bekommen, übernehmen wir das folgende Theorem 2.11 aus ALEXANDER(1984a):

2.1.6 Theorem. Für alle $n \in \mathbb{N}$ und $\epsilon \geq 8n^{-1/2}$ gilt:

$$\mathbb{P}(\sup\{|\mu_n(C) - \nu_n(C)| : C \in \mathcal{C}\} \geq \epsilon) \leq 16(n\epsilon^2)^{2^{11}V(C)} \exp(-2n\epsilon^2).$$

2.1.7 Korollar. Für alle $0 < s < 2$ existiert eine positive Konstante K , die nur von s und $\mathcal{C}$ abhängt, so daß für $n \in \mathbb{N}$ gilt:

$$\mathbb{E}(\exp((2-s)nD_n^2)) \leq K, \text{ wobei } D_n := \sup\{|\mu_n(C) - \nu_n(C)| : C \in \mathcal{C}\}.$$

Beweis. Sei $0 < s < 2$ fest vorgegeben. Da für $\epsilon \geq \exp(64(2-s))$

gilt: $(\frac{\log \epsilon}{(2-s)n})^{1/2} \geq 8n^{-1/2}$, erhalten wir für alle $n \in \mathbb{N}$ wegen

2.1.6:

$$\begin{aligned} \mathbb{E}(\exp((2-s)nD_n^2)) &= 1 + \int_1^\infty \mathbb{P}(D_n \geq (\frac{\log \epsilon}{(2-s)n})^{1/2}) d\epsilon \\ &\leq \exp(64(2-s)) + \int_1^\infty \min\{1, h(s, \epsilon)\} d\epsilon, \end{aligned}$$

$$\text{wobei } h(s, \epsilon) := 16 \cdot (\frac{\log \epsilon}{2-s})^{2^{11}V(C)} \exp(-\frac{2}{2-s} \log \epsilon)$$

$$= 16 \cdot (2-s)^{-2^{11}V(C)} (\log \epsilon)^{2^{11}V(C)} \epsilon^{-\frac{2}{2-s}},$$

d.h. wegen $\frac{2}{2-s} > 1$ folgt sofort die Behauptung. $\square$

2.2 Relative L_b -Folgenkompaktheit, L_b -Konvergenz sowie Glivenko-Cantelli-Aussagen für Dreiecksschemata

Wir wenden uns nun wieder der Situation zu, die wir eingangs von Kapitel II betrachtet haben (es muß also nicht mehr gelten: $i(n) = n$ für alle $n \in \mathbb{N}$). Dann folgt zunächst:

2.2.1 Theorem (Glivenko-Cantelli). Sei $(\gamma_n)_{n \in \mathbb{N}} \subset (1, \infty)$ mit

$\sum_{n=1}^{\infty} \gamma_n^{-\varepsilon} < \infty$ für alle $\varepsilon > 0$. Dann gilt :

$$(\log \gamma_n)^{-1/2} \sup\{|\beta_n(C)| : C \in \mathcal{C}^-\} \rightarrow 0 \text{ } \mathbb{P}\text{-f.s.}$$

Beweis. Dies ist eine unmittelbare Folgerung aus 2.1.7 (angewandt auf $\mathcal{C}^-$ für $s=1$), denn für alle $n \in \mathbb{N}$ und $\varepsilon > 0$ erhalten wir durch Anwendung der Markoffungleichung (mit

$$M_n := \sup\{|\beta_n(C)| : C \in \mathcal{C}^-\} :$$

$$\mathbb{P}((\log \gamma_n)^{-1/2} M_n \geq \varepsilon) \leq \mathbb{P}(\exp(M_n^2) \geq \gamma_n^{\varepsilon^2})$$

$$\leq \gamma_n^{-\varepsilon^2} \cdot \mathbb{E}(\exp(M_n^2))$$

$$\leq K \gamma_n^{-\varepsilon^2}.$$

Dabei ist K lediglich von $\mathcal{C}^-$ abhängig, d.h. aus der vorausgesetzten Summierbarkeit folgt nun gemäß Borel-Cantelli die Behauptung. □

2.2.2 Bemerkung.

(i) Aus Theorem 2.2.1 folgt insbesondere (man vgl. hierzu auch entsprechende Ergebnisse in GÄNSSLER(1985), die dort Folgerungen sind aus der Vapnik-Chervonenkis-Ungleichung im "i.i.d.-Fall):

$$K_n (\log n)^{-1/2} M_n \rightarrow 0 \text{ } \mathbb{P}\text{-f.s.}, \text{ falls } K_n \rightarrow 0 \text{ für } n \rightarrow \infty$$

oder

$$i(n)^{-\gamma} M_n \rightarrow 0 \text{ } \mathbb{P}\text{-f.s. für alle } \gamma > 0, \text{ falls } \sum_{n=1}^{\infty} i(n)^{-K} < \infty$$

für ein $K > 0$.

Dabei ist M_n definiert wie im Beweis zu 2.2.1.

(ii) Ohne eine zusätzliche Voraussetzung an die gemeinsame Verteilung der Zeilen des Dreiecksschemas lassen sich diese Ergebnisse i.a. nicht verschärfen.

Wir wollen nun die relative L_b -Folgenkompaktheit von $(\beta_n)_{n \in \mathbb{N}}$ untersuchen. Um dies im Rahmen der Ergebnisse von Kapitel I tun zu können, betrachten wir für jedes $n \in \mathbb{N}$ die Halbmetrik

$$s_n(C, D) := \mathbb{E}([\beta_n(C) - \beta_n(D)]^2)^{1/2} \text{ für } C, D \in \mathcal{C}$$

auf $\mathcal{C}$ und fordern:

(A) Es existiere eine Halbmetrik s auf $\mathcal{C}$ mit:

$$(i) \lim_{n \rightarrow \infty} \sup \sup \{ s_n(C, D) : s(C, D) \leq \delta; C, D \in \mathcal{C} \} = 0 \text{ für } \delta > 0.$$

(ii) $(\mathcal{C}, s)$ ist totalbeschränkt.

2.2.3 Bemerkung. Für alle $C, D \in \mathcal{C}$ und $n \in \mathbb{N}$ gilt natürlich:

$s_n(C, D) \leq 2^{1/2} \cdot v_n(CAD)^{1/2}$. (A) ist also insbesondere erfüllt, falls ein finites Maß v auf $(Y, \mathcal{B})$ existiert mit:

$$\lim_{n \rightarrow \infty} \sup \sup \{ v_n(C) : C \in \mathcal{C}^-; v(C) \leq \delta \} = 0 \text{ für } \delta > 0 \text{ (als geeignete}$$

Halbmetrik wählt man $s(C, D) := v(CAD)$ für $C, D \in \mathcal{C}$, denn mit Lemma (7.13) in DUDLEY(1978) ergibt sich sofort die Totalbeschränktheit von $(\mathcal{C}, s)$). Die Bedingung (A) ist ähnlich zur Forderung (3.9) in ALEXANDER(1984b).

Die in Theorem 2.2.4 bzw. 2.2.5 auftretenden " L_b -Aussagen" werden stets unter der Gültigkeit von (A) hergeleitet und dann im folgenden Modell formuliert:

Sei $X := \mathcal{C}$, s die Halbmetrik aus (A), $S := \ell^\infty(X)$,

$A := \sigma(\pi_C(\mathcal{C}) : C \in \mathcal{C})$ mit $\pi_C(\mathcal{C})(f) := f(C)$ für $f \in \ell^\infty(X)$.

Dann können die $(\beta_n)_{n \in \mathbb{N}}$ als F, A -meßbare Zufallselemente in $\ell^\infty(X)$ aufgefaßt werden, d.h. die Ausgangssituation von Kapitel I wäre gesichert. Durch die Definition von A sind überdies auch die Bedingungen (M2) bzw. (M3) erfüllt, wohingegen (M1) eine unmittelbare Konsequenz ist aus der vorausgesetzten Abzählbarkeit von C .

2.2.4 Theorem. Falls die Bedingung (A) erfüllt ist und $i(n) \rightarrow \infty$ für $n \rightarrow \infty$, folgt: $(\beta_n)_{n \in \mathbb{N}}$ ist relativ L_b -folgenkompakt.

Beweis. Für jedes $n \in \mathbb{N}$ und $\delta > 0$ definieren wir:

$$\omega_{\beta_n}^r(\delta) := \sup\{|\beta_n(C) - \beta_n(D)| : r(C, D) \leq \delta; C, D \in C\} \text{ für } r \in \{s, s_n\},$$

$$a_n(\delta) := \sup\{s_n(C, D) : s(C, D) \leq \delta; C, D \in C\} \text{ sowie ein}$$

$$N_n(\delta) \subset C \text{ mit den Eigenschaften:}$$

(i) Für alle $C \in C$ existiert ein $A(n, C) \in N_n(\delta)$ mit

$$v_n(C \Delta D) \leq \delta.$$

(ii) $|N_n(\delta)|$ ist minimal unter Bedingung (i).

Wegen (VC3) folgt natürlich: $|N_n(\delta)| \leq K\delta^{-1-V(C)}$, wobei K lediglich von C abhängt.

Weiter gilt (da $s_n(C, D) \leq [2v_n(C \Delta D)]^{\frac{1}{2}}$) für alle $n \in \mathbb{N}$, $0 < \delta \leq 1$:

$$\begin{aligned} \omega_{\beta_n}^{s_n}(\delta) &\leq 2 \sup\{|\beta_n(C) - \beta_n(A(n, C))| : C \in C\} \\ &\quad + \sup\{|\beta_n(C) - \beta_n(D)| : C, D \in N_n(\delta); s_n(C, D) \leq \delta + 3\delta^{1/2}\} \\ &\leq 2 \tilde{\omega}_{\beta_n}(\delta) + R_n(\delta), \text{ wobei} \end{aligned}$$

$\tilde{\omega}_{\beta_n}(\delta)$ definiert ist wie in 2.1.4(i) und

$$R_n(\delta) := \sup\{|\beta_n(C) - \beta_n(D)| : C, D \in N_n(\delta); s_n(C, D) \leq 4\delta^{1/2}\}.$$

Mit Hilfe der Bernsteinungleichung erhalten wir daraus:

$$\begin{aligned} (*) \quad \lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(\omega_{\beta_n}^{S_n}(\delta) \geq \epsilon) &\stackrel{2.1.4(i)}{\leq} \\ &\leq \limsup_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(R_n(\delta) \geq \frac{\epsilon}{2}) \\ &\leq 2K^2 \lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \delta^{-2-2V(C)} \exp\left(-\frac{\epsilon^2/4}{32\delta + \epsilon i(n)^{-1/2}}\right) \\ &= 0 \quad \text{für alle } \epsilon > 0. \end{aligned}$$

Unter Berücksichtigung von (A)(i) liefert dies:

$$\begin{aligned} \lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(\omega_{\beta_n}^S(\delta) \geq \epsilon) &\leq \lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(\omega_{\beta_n}^{S_n}(a_n(\delta)) \geq \epsilon) \\ &\leq \limsup_{n \rightarrow \infty} \mathbb{P}(\omega_{\beta_n}^{S_n}(\tilde{\delta}) \geq \epsilon) \quad \text{für beliebige } \epsilon, \tilde{\delta} > 0. \end{aligned}$$

Mit $\tilde{\delta} \downarrow 0$ folgt nun unter Beachtung von (*), daß die Bedingung (iii)(A) aus 1.3.4 erfüllt ist. Die Bedingung (iii)(B) hingegen bereitet keine Schwierigkeiten, denn aus der Tschebyscheffungleichung ergibt sich für alle $C \in \mathcal{C}$:

$$\limsup_{n \rightarrow \infty} \mathbb{P}(|\beta_n(C)| \geq M) \leq M^{-2} \rightarrow 0 \quad \text{für } M \rightarrow \infty.$$

Aus 1.3.4 folgt also die Behauptung. $\square$

Mit Hilfe des klassischen Zentralen Grenzwertsatzes und 1.4.1 erhalten wir (entspricht Theorem 3.1 in ALEXANDER(1985)):

2.2.5 Theorem. Es gelte (A) sowie $i(n) \rightarrow \infty$ für $n \rightarrow \infty$. Dann sind äquivalent:

- (i) $\beta_n \xrightarrow{L_b} G$, wobei G ein zentrierter Gaußprozeß ist.
- (ii) Die Kovarianzfunktion von $(\beta_n)_{n \in \mathbb{N}}$ konvergiert, d.h. für alle $C, D \in \mathcal{C}$ gilt: $(\mathbb{E}(\beta_n(C)\beta_n(D)))_{n \in \mathbb{N}}$ ist konvergent.

2.2.6 Bemerkung.

(i) Ein erster "Kandidat" für s in (A) wird natürlich stets $s(C,D) := \limsup_{n \rightarrow \infty} v_n(C \Delta D)$ sein (es ist unmittelbar klar, daß durch s eine Halbmetrik auf $\mathcal{C}$ definiert ist). Sofern man dafür die Bedingungen (A)(i) bzw. (A)(ii) nachweisen kann, sind die Theoreme 2.2.4 bzw. 2.2.5 anwendbar. Man vergleiche in diesem Zusammenhang das Theorem 3.4 in GÄNSSLER-HÄUSLER-SCHNEEMEIER(1984), wo in der Situation $\mathcal{C} = \{[0,t] : t \in [0,1]\}$ die Bedingung (3.6) auftauchte mit:

- (*) Es existiert eine monoton wachsende und stetige Funktion $G: [0,1] \rightarrow \mathbb{R}$ mit $\limsup_{n \rightarrow \infty} v_n([0,r] \Delta [0,t]) \leq G(t) - G(r)$ für $0 \leq r \leq t \leq 1$.

Wenn wir nun definieren: $s([0,r], [0,t]) := |t - r|$ für $r, t \in [0,1]$, folgt natürlich:

$$\sup\{v_n(C \Delta D) : s(C,D) \leq \frac{1}{2K} ; C, D \in \mathcal{C}\} \leq \sup\{v_n\left(\left[\frac{1-2}{2K}, \frac{1}{2K}\right] : i=2,3,\dots,2K\right) \\ \text{für alle } K \in \mathbb{N}, \text{ also}$$

$$\limsup_{n \rightarrow \infty} \sup\{v_n(C \Delta D) : s(C,D) \leq \frac{1}{2K} ; C, D \in \mathcal{C}\} \leq \\ \leq \sup\{G\left(\frac{1}{2K}\right) - G\left(\frac{1-2}{2K}\right) : i=2,3,\dots,2K\} \rightarrow 0 \text{ für } K \rightarrow \infty.$$

Da wegen (VC3) auch (A)(ii) erfüllt ist, ist also (A) gezeigt.

(ii) Wie man aus dem Beweis zu (i) erkennt, ist (A) immer erfüllt, wenn gilt: Es existiert ein finites Maß ν auf $(Y, \mathcal{B})$ mit:

$$(1) \lim_{n \rightarrow \infty} \sup \nu_n(C) \leq \nu(C) \text{ für alle } C \in \mathcal{C}^-.$$

(2) Für alle $\delta > 0$ gibt es $C_1, \dots, C_k \in \mathcal{B}$ mit der Eigenschaft:
für $C \in \mathcal{C}$ existieren $i, j=1, \dots, k$ mit $C_1 \subset C \subset C_j$ und
 $\nu(C_j \setminus C_1) \leq \delta$.

2.3 Der gebootstrappte empirische C-Prozeß

Wir gehen wieder von der Situation aus, die wir zu Beginn von Kapitel II betrachtet haben. Sei weiter $(\epsilon_{ni})_{i \leq m(n), n \in \mathbb{N}}$ ein von $(\xi_{ni})_{i \leq i(n), n \in \mathbb{N}}$ unabhängiges Dreiecksschema von zeilenweise unabhängigen Variablen über $(\Omega, \mathcal{F}, \mathbb{P})$, wobei für jedes $n \in \mathbb{N}$ und $i \leq m(n)$ ϵ_{ni} gleichverteilt sei über $\{1, \dots, i(n)\}$. Dann nennen wir

$$B_n(C) := m(n)^{-1/2} \sum_{i=1}^{m(n)} [1_{C(\xi_{n\epsilon_{ni}})} - \mu_n(C)], C \in \mathcal{C}$$

den gebootstrappten empirischen C-Prozeß zu $\xi_{n1}, \dots, \xi_{ni(n)}$ (bzl. eines "Datenmultiplikators" $m(n)$).

2.3.1 Bemerkung. Diese Definition ist selbstverständlich nur eindeutig bis auf Verteilung, da es davon abhängt, wie man die Daten produziert (d.h. welche $(\epsilon_{ni})_{i \leq m(n), n \in \mathbb{N}}$ man wählt).

Aus der Definition erhalten wir unmittelbar:

2.3.2 Lemma. Für alle $n \in \mathbb{N}$ und $C, D \in \mathcal{C}$ gilt:

(i) $\mathbb{E}(B_n(C)) = 0.$

(ii) $\mathbb{E}(B_n(C)B_n(D) \mid \xi_{n1}, \dots, \xi_{ni(n)}) \stackrel{[\mathbb{P}]}{=} \mu_n(C \cap D) - \mu_n(C)\mu_n(D).$

(iii) Für $Q(\xi_{n1}, \dots, \xi_{ni(n)})$ -f.a. $(x_1, \dots, x_{i(n)}) \in Y^n$ gilt:

$$L((\xi_{n\epsilon_{n1}}, \dots, \xi_{n\epsilon_{nm(n)}}) \mid (\xi_{n1}, \dots, \xi_{ni(n)}) = (x_1, \dots, x_{i(n)})) =$$

$$L((x_{\epsilon_{n1}}, \dots, x_{\epsilon_{nm(n)}})), \text{ wobei die } x_{\epsilon_{ni}}, i=1, \dots, m(n),$$

unabhängig und identisch verteilt sind mit:

$$\mathbb{P}(x_{\epsilon_{n1}} = x_j) = \frac{|\{k \leq i(n) : x_k = x_j\}|}{i(n)} = i(n)^{-1} \sum_{k=1}^{i(n)} 1_{\{x_j\}}(x_k)$$

für alle $j=1, \dots, i(n).$

Beweis. Ergibt sich aus GÄNSSLER-STUTE(1977), 5.3.22.

Wir betrachten nun für festes $\bar{x} = (x_{ni})_{i \leq i(n), n \in \mathbb{N}}, x_{ni} \in Y,$

den stochastischen Prozeß $B_n^{\bar{x}} = (B_n^{\bar{x}}(C))_{C \in \mathcal{C}}$, definiert

durch

$$B_n^{\bar{x}}(C) := m(n)^{-1/2} \sum_{i=1}^{m(n)} [1_C(x_{n\epsilon_{ni}}) - \mu_n^{\bar{x}}(C)], \text{ wobei}$$

$$\mu_n^{\bar{x}}(C) := i(n)^{-1/2} \sum_{i=1}^{i(n)} 1_C(x_{ni}) \text{ mit } C \in \mathcal{C}, n \in \mathbb{N}.$$

Gemäß 2.3.2(iii) ist also $B_n^{\bar{x}}$ der empirische C -Prozeß zu

$x_{n\epsilon_{n1}}, \dots, x_{n\epsilon_{nm(n)}}$, so daß wir die in Abschnitt 2.2 bereit-

gestellten Ergebnisse anwenden können. Dazu betrachten wir

die folgende Forderung (B) (ein Analogon zur Bedingung (A)

aus 2.2):

(B) Es existiert eine Halbmetrik s auf C mit:

$$(i) \lim_{n \rightarrow \infty} \sup \sup \{ v_n(C \setminus D) (1 - v_n(C \setminus D)) : s(C, D) \leq \delta; C, D \in C \} + 0$$

für $\delta > 0$.

(ii) (C, s) ist totalbeschränkt.

Die folgenden L_b -Aussagen werden wir nun stets unter der Gültigkeit von (B) herleiten und dann im Modell von Abschnitt 2.2 formulieren (mit s aus (B); man beachte, daß die $x_{n \in ni}$, $i = 1, \dots, m(n)$, trivialerweise $F, \mathcal{B}$ -meßbar sind).

2.3.3 Theorem (Glivenko-Cantelli). Für $(\gamma_n)_{n \in \mathbb{N}} \subset (1, \infty)$ mit

$$\sum_{n=1}^{\infty} \gamma_n^{-\varepsilon} < \infty \text{ für alle } \varepsilon > 0 \text{ folgt:}$$

$$(\log \gamma_n)^{-1/2} \sup \{ |B_n^{\bar{x}}(C)| : C \in C \} \rightarrow 0 \text{ } \mathbb{P}\text{-f.s. für alle}$$

$$\bar{x} = (x_{ni})_{i \leq i(n), n \in \mathbb{N}}, x_{ni} \in Y.$$

Beweis. Ergibt sich unmittelbar aus Theorem 2.2.1.

Mit $\bar{\xi} := (\xi_{ni})_{i \leq i(n), n \in \mathbb{N}}$ erhalten wir weiter:

2.3.4 Theorem. Sei (B) erfüllt, $\sum_{n=1}^{\infty} \exp(-\varepsilon i(n)) < \infty$ für alle

$\varepsilon > 0$ sowie $m(n) \rightarrow \infty$ für $n \rightarrow \infty$. Für $Q_{\bar{\xi}}$ -fast alle $\bar{x}$ gilt dann:

$(B_n^{\bar{x}})_{n \in \mathbb{N}}$ ist relativ L_b -folgenkompakt, wobei mögliche Häufungspunkte von $(L(B_n^{\bar{x}}))_{n \in \mathbb{N}}$ nicht von $\bar{x}$ abhängen.

Beweis. Die vorausgesetzte Reihenbedingung liefert zusammen mit 2.2.1, daß für $Q_{\bar{\xi}}$ -fast alle $\bar{x}$ gilt:

$$(*) \sup \{ |\mu_n^{\bar{x}}(C) - v_n(C)| : C \in C^- \} \rightarrow 0 \text{ für } n \rightarrow \infty.$$

Hiermit ist aber insbesondere für Q_{ξ} -fast alle $\bar{x}$ die folgende Bedingung erfüllt:

($A_{\bar{x}}$) Es existiert eine Halbmetrik s auf C (ebendie aus (B))

mit:

$$(i) \lim_{n \rightarrow \infty} \sup \sup \{ \mathbb{E}((B_n^{\bar{x}}(C) - B_n^{\bar{x}}(D))^2)^{1/2} : s(C, D) \leq \delta; C, D \in C \} \rightarrow 0 \text{ für } \delta \rightarrow 0.$$

(ii) (C, s) ist totalbeschränkt.

Wegen ($A_{\bar{x}}$) und $m(n) \rightarrow \infty$ für $n \rightarrow \infty$ ergibt sich nun die Behauptung sofort aus Theorem 2.2.4, nachdem jeder mögliche Häufungspunkt von $(L(B_n^{\bar{x}}))_{n \in \mathbb{N}}$ gemäß (*) nur von $(v_n)_{n \in \mathbb{N}}$ abhängt. $\square$

2.3.5 Theorem. Unter den Voraussetzungen von 2.3.4 sind für Q_{ξ} -fast alle $\bar{x}$ die beiden folgenden Aussagen äquivalent:

(i) Es existiert ein von $\bar{x}$ unabhängiger zentrierter Gaußprozeß $G = (G(C))_{C \in C}$ mit Kovarianzstruktur $K(C, D) = \mathbb{E}(G(C)G(D))$ für $C, D \in C$ und $B_n^{\bar{x}} \xrightarrow{L_b} G$.

(ii) Für alle $C, D \in C$ gilt: $v_n(C \cap D) - v_n(C)v_n(D) \rightarrow K(C, D)$ für $n \rightarrow \infty$.

Beweis. Aus 2.3.4 und dem Zentralen Grenzwertsatz für Dreieckschemata folgt sofort (analog zum Beweis von Theorem 18 (iii) in GÄNSSLER(1983)) die Äquivalenz von (i) mit:

(ii') $\mu_n^{\bar{x}}(C \cap D) - \mu_n^{\bar{x}}(C)\mu_n^{\bar{x}}(D) \rightarrow K(C, D)$ für $n \rightarrow \infty$ und $C, D \in C$.

Wegen (*) im Beweis zu 2.3.4 (Gültigkeit von Glivenko-Cantelli) folgt daraus für Q_{ξ} -fast alle $\bar{x}$ sofort die Äquivalenz von (ii') mit (ii) und damit die Behauptung. $\square$

2.3.6 Bemerkung. Aus Theorem 2.3.5 folgt insbesondere das Hauptergebnis von GÄNSSLER(1985).

Nachdem wir nun für $Q_{\bar{\xi}}$ -fast alle $\bar{x}$ Aussagen über $(B_n^{\bar{x}})_{n \in \mathbb{N}}$ hergeleitet haben, stellt sich natürlich die Frage, ob diese in entsprechender Form auch für $(B_n)_{n \in \mathbb{N}}$ gelten. Die nun folgenden L_b -Aussagen werden wir wieder herleiten unter der Bedingung (B) und formulieren im Modell von 2.2. Durch Integration über die Realisierungen $\bar{x}$ von $\bar{\xi}$ erhalten wir dann die folgenden Theoreme:

2.3.7 Theorem. Für $(\gamma_n)_{n \in \mathbb{N}} \subset (1, \infty)$ mit $\sum_{n=1}^{\infty} \gamma_n^{-\epsilon} < \infty$ für alle $\epsilon > 0$ gilt:

$$(\log \gamma_n)^{-1/2} \sup\{|B_n(C)| : C \in \mathcal{C}\} \longrightarrow 0 \text{ } \mathbb{P}\text{-f.s.}$$

Beweis. Aus 2.3.2(iii) und 2.3.3 folgt unmittelbar:

$$\mathbb{P}(\lim_{n \rightarrow \infty} (\log \gamma_n)^{-1/2} \sup\{|B_n(C)| : C \in \mathcal{C}\} = 0) =$$

$$\int \mathbb{P}(\lim_{n \rightarrow \infty} (\log \gamma_n)^{-1/2} \sup\{|B_n^{\bar{x}}(C)| : C \in \mathcal{C}\} = 0) Q_{\bar{\xi}}(d\bar{x}) = 1. \quad \square$$

2.3.8 Theorem. Sei (B) erfüllt, $\sum_{n=1}^{\infty} \exp(-\epsilon i(n)) < \infty$ für alle

$\epsilon > 0$ sowie $m(n) \rightarrow \infty$ für $n \rightarrow \infty$. Dann ist $(B_n)_{n \in \mathbb{N}}$ relativ L_b -folgenkompakt.

Beweis. Für beliebiges $\epsilon > 0$ gilt (wobei $\omega_F(\cdot)$ wie in 1.1.1 definiert ist mit $X=C$ und s aus (B)):

$$\lim_{\delta \downarrow 0} \lim_{n \rightarrow \infty} \sup \mathbb{P}(\omega_{B_n}(\delta) \geq \epsilon) \stackrel{2.3.2(iii)}{=} 0$$

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \int \mathbb{P}(\omega_{B_n^X}(\delta) \geq \varepsilon) Q_{\xi}(d\bar{X}) \stackrel{\text{Lemma von Fatou}}{\leq}$$

$$\int \lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(\omega_{B_n^X}(\delta) \geq \varepsilon) Q_{\xi}(d\bar{X}) \stackrel{2.3.4}{=} 0.$$

$$\text{Analog folgt für alle } C \in \mathcal{C}: \lim_{K \uparrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(|B_n(C)| \geq K) = 0.$$

Mit 1.3.4 ergibt sich daraus die Behauptung. $\square$

Auf ähnliche Weise ergibt sich aus 2.3.5:

2.3.9 Theorem. Unter den Voraussetzungen von 2.3.8 folgt die Äquivalenz der beiden folgenden Aussagen:

(i) Es existiert ein zentrierter Gaußprozeß $G = (G(C))_{C \in \mathcal{C}}$

mit Kovarianzstruktur K und $B_n \xrightarrow{L_b} G$.

(ii) Für alle $C, D \in \mathcal{C}$ gilt: $v_n(C \cap D) - v_n(C) v_n(D) \rightarrow K(C, D)$

für $n \rightarrow \infty$.

2.3.10 Bemerkung. Das "Bootstrappen" kann in zweierlei Hinsicht von Bedeutung sein:

(i) Da im Normalfall $(v_n)_{n \in \mathbb{N}}$ unbekannt sein dürfte, kann der Statistiker natürlich auch $(\beta_n)_{n \in \mathbb{N}}$ nicht bestimmen. In diesem Fall bietet es sich an, $(B_n)_{n \in \mathbb{N}}$ zu verwenden, da es allein durch den Datensatz und dessen "Vervielfachen" (das "Bootstrappen") bestimmt ist.

(ii) Vom beweistechnischen Standpunkt aus kann dem Bootstrap-Verfahren auch große Bedeutung zukommen, denn es gilt

im Fall $i(n) = m(n)$: $\beta_n = -B_n + \gamma_n$ für alle $n \in \mathbb{N}$, wobei

γ_n den empirischen C -Prozeß bezeichne zu $\xi_{n \in n_1}, \dots, \xi_{n \in n_m(n)}$.

Dabei liegt offensichtlich im Fall $(\gamma_n)_{n \in \mathbb{N}}$ wie auch im Fall $((B_n)_{n \in \mathbb{N}}$ unter der Bedingung $\bar{\xi} = \bar{x}$ bei festem $\bar{x}$) jeweils ein zeilenweise unabhängiges, zeilenweise identisch verteiltes Dreiecksschema zugrunde, d.h.: "Bootstrappen stellt eine Möglichkeit dar, den non-i.i.d.-Fall in den i.i.d.-Fall überzuführen."

2.4 Bemerkungen zum Fall gewichteter empirischer C-Prozesse

Wir wollen wieder von der Situation ausgehen, die wir zu Beginn von Kapitel II betrachtet haben. Sei weiter

$$\begin{aligned} q: [0,1] \rightarrow \mathbb{R} \quad \text{eine Funktion mit: } q(0) &= 0; \quad q \text{ ist stetig und} \\ &\text{strikt wachsend auf } [0, \frac{1}{2}); \\ q \text{ ist symmetrisch um } \frac{1}{2}; \\ \int_0^1 q^{-2}(x) dx &< \infty. \end{aligned}$$

Dann definieren wir für jedes $n \in \mathbb{N}$ und $C \in \mathcal{C}$:

$$\bar{w}_n(C) := \begin{cases} q^{-1}(\mathbb{E}(\beta_n(C)^2)) \beta_n(C) & \text{falls } \mathbb{E}(\beta_n(C)^2) > 0, \\ 0 & \text{falls } \mathbb{E}(\beta_n(C)^2) = 0. \end{cases}$$

Es ist nun unmittelbar klar, daß dieser gewichtete C-Prozeß $\bar{w}_n$ i.a. kein Element ist aus $\ell^\infty(C)$. Um also L_b -Aussagen im Modell von Kapitel I formulieren zu können, müssen wir den folgenden Prozeß betrachten für $C \in \mathcal{C}$:

$$w_n(C) := \begin{cases} \bar{w}_n(C) & \text{falls } \mathbb{E}(\beta_n(C)^2) \geq (2i(n))^{-1} \log i(n), \\ 0 & \text{falls } \mathbb{E}(\beta_n(C)^2) < (2i(n))^{-1} \log i(n). \end{cases}$$

Diese Prozesse werden in ALEXANDER(1985) untersucht. Mit Hilfe des dortigen Theorems 5.1 (eine zentrale stochastische

Ungleichung) können wir das folgende Resultat formulieren (wobei wir wieder das Modell von 2.2 zugrundelegen):

2.4.1 Theorem. Falls (A) aus 2.2 erfüllt ist sowie $i(n) \rightarrow \infty$ für $n \rightarrow \infty$, folgt: $(W_n)_{n \in \mathbb{N}}$ ist relativ L_b -folgenkompakt.

In gleicher Weise erhalten wir dann ein entsprechendes Resultat zu 2.2.5. Auf diesen Ergebnissen aufbauend ergeben sich nun wie in 2.3 gewichtete "Bootstrap"-Aussagen. Unter Verwendung der Definitionen des Abschnitts 2.3 betrachten wir für alle $\bar{x} = (x_{ni})_{i \leq n, n \in \mathbb{N}}$, $C \in \mathcal{C}$ sowie $n \in \mathbb{N}$:

$$\bar{B}_n^{\bar{x}}(C) := \begin{cases} q^{-1}(\mu_n^{\bar{x}}(C)) \cdot B_n^{\bar{x}}(C) & \text{falls } \mu_n^{\bar{x}}(C) \in [b_n, 1-b_n], \\ 0 & \text{sonst.} \end{cases}$$

$$\bar{B}_n(C) := \begin{cases} q^{-1}(\mu_n(C)) \cdot B_n(C) & \text{falls } \mu_n(C) \in [b_n, 1-b_n], \\ 0 & \text{sonst.} \end{cases}$$

Dabei ist $b_n := m(n)^{-1} \log m(n)$.

Dann folgt, wobei wir wieder das Modell von 2.2 verwenden:

2.4.2 Theorem. Falls (B) aus 2.3 erfüllt ist sowie

$\sum_{n=1}^{\infty} \exp(-\varepsilon i(n)) < \infty$ für alle $\varepsilon > 0$ und $m(n) \rightarrow \infty$ für $n \rightarrow \infty$, folgt:

- (i) $(\bar{B}_n^{\bar{x}})_{n \in \mathbb{N}}$ ist relativ L_b -folgenkompakt für Q_{ξ} -fast alle $\bar{x}$.
- (ii) $(\bar{B}_n)_{n \in \mathbb{N}}$ ist relativ L_b -folgenkompakt.

2.4.3 Bemerkung. Natürlich können auf der Basis von 2.4.2

nun auch Aussagen zur L_b -Konvergenz hergeleitet werden (analog zu 2.3.5 bzw. 2.3.9).

KAPITEL III: L_b -AUSSAGEN UND GLIVENKO-CANTELLI-RESULTATE
FÜR U-PROZESSE, DIE INDIZIERT SIND NACH VAP-
NIK-CHERVONENKIS-KLASSEN VON MENGEN

3.1 Definition des U-Prozesses und elementare Eigenschaften

Sei wieder $(\xi_{ni})_{i \leq i(n), n \in \mathbb{N}}$ ein Dreiecksschema von zeilenweise unabhängigen (nicht notwendig identisch verteilten) Zufallselementen in einen meßbaren Raum $(Z, \mathcal{D})$, definiert über einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ und $\mathcal{F}, \mathcal{D}$ -meßbar. Sei weiter für den gesamten Verlauf von Kapitel III m eine feste natürliche Zahl (o.E. $i(n) \geq m$ für alle $n \in \mathbb{N}$) sowie $C \subset \bigotimes_{i=1}^m \mathcal{D}$ eine abzählbare Vapnik-Chervonenkis-Klasse. Für jedes $n \in \mathbb{N}$ definieren wir dann:

3.1.1 Definition. Sei $D \in \bigotimes_{i=1}^m \mathcal{D}$.

$$E_n := \{(i_1, \dots, i_m) \in \{1, \dots, i(n)\}^m : i_j \neq i_k \text{ für } j \neq k \text{ mit } j, k = 1, \dots, m\},$$

$$H_n(D) := |E_n|^{-1} \sum_{(i_1, \dots, i_m) \in E_n} 1_D(\xi_{ni_1}, \dots, \xi_{ni_m}),$$

$$\bar{H}_n(D) := \mathbb{E}(H_n(D)),$$

$$U_n(D) := i(n)^{1/2} [H_n(D) - \bar{H}_n(D)].$$

Dann heißt $U_n = (U_n(C))_{C \in C}$ für jedes $n \in \mathbb{N}$ der nach C indizierte U-Prozeß (zu $\xi_{n1}, \dots, \xi_{ni(n)}$), wobei klar ist:

$$|E_n| = i(n) \cdot (i(n)-1) \cdot \dots \cdot (i(n)-m+1).$$

3.1.2 Bemerkung.

(i) Mit dieser Bezeichnung lehnen wir uns an NOLAN-POLLARD(1985)

an, wo Glivenko-Cantelli-Konvergenzraten untersucht werden. Allerdings wird dort nicht die Situation von Dreiecksschemata betrachtet und auch nur die i.i.d.-Situation für den Fall $m=2$ abgehandelt.

(ii) Ursprünglich geht die Untersuchung dieser Prozesse zurück auf SERFLING(1984), der die folgende Situation betrachtet:

Für eine $\mathcal{B}_m, \mathcal{B}^*$ -meßbare "Kernfunktion" $h: \mathbb{R}^m \rightarrow \mathbb{R}$ sowie eine Folge $(X_i)_{i \in \mathbb{N}}$ von unabhängigen und identisch verteilten zufälligen Variablen betrachtet man mit $n \in \mathbb{N}$ die der Größe nach geordneten Werte $W_1^n \leq \dots \leq W_{r_n}^n$ ($r_n = |E_n|$) von $h(X_{i_1}, \dots, X_{i_m})$, $(i_1, \dots, i_m) \in E_n - E_n$ wie in 3.1.1 mit $i(n)=n$. Man untersucht nun die "verallgemeinerte l -Statistik"

$$T_n := r_n^{-1} \sum_{i=1}^{r_n} c_{ni} W_i^n,$$

wobei $(c_{ni})_{i \leq r_n, n \in \mathbb{N}}$ ein fest vorgegebenes Dreiecksschema von Konstanten ("Scores") ist. Wenn wir nun das Schema

$(\xi_{ni})_{i \leq n, n \in \mathbb{N}}$ betrachten mit $\xi_{ni} := X_i$ für $i \leq n, n \in \mathbb{N}$ sowie die Vapnik-Chervonenkis-Klasse $C = \{h^{-1}((-\infty, t]): t \in \mathbb{R}\}$, so folgt mit den Definitionen aus 3.1.1 (H_n^{-1} ist Quantiltransformierte von H_n):

$$T_n = \int_0^1 J_n(t) H_n^{-1}(t) dt, \text{ wobei für } i=1, \dots, r_n \text{ gilt:}$$

$$J_n(t) := c_{ni} \text{ für } \frac{i-1}{r_n} < t \leq \frac{i}{r_n}.$$

Die Untersuchung von Glivenko-Cantelli-Verhalten bzw. Verteilungskonvergenz für $(T_n)_{n \in \mathbb{N}}$ läßt sich also unter geeigneten Voraussetzungen an $(J_n)_{n \in \mathbb{N}}$ zurückführen auf entsprechende Ergebnisse für den U-Prozeß (man beachte, daß die Abzählbarkeitsbedingung an C keine Einschränkung darstellt).

(iii) Schwache Konvergenz für gewichtete U-Prozesse in der Situation $\{(X_i)_{i \in \mathbb{N}} \text{ sind i.i.d.}; C = \{h^{-1}((-\infty, t]): t \in \mathbb{R}\} \text{ für eine Kernfunktion } h\}$ wird bei SILVERMAN(1983) untersucht. Hierauf wollen wir jedoch erst in Kapitel IV eingehen, wo wir unter Verwendung einfacherer Methoden eine Verallgemeinerung des dortigen Hauptergebnisses erzielen werden (bei Silverman wird KMT benutzt).

(iv) Es ist unmittelbar klar, daß U_n für jedes $n \in \mathbb{N}$ ein Zufallselement ist in $(\ell^\infty(C), A)$, wobei wiederum $A := \sigma(\pi_C(C) : C \in C)$ mit $\pi_C(C)(f) := f(C), f \in \ell^\infty(C)$. Wir werden also wieder in der Lage sein, die Ergebnisse im Rahmen des Modells von Kapitel I zu formulieren. Darüberhinaus ist klar, daß die Abzählbarkeitsvoraussetzung an C keine wesentliche Beschränkung darstellt und i.a. jederzeit durch geeignete Meßbarkeitsforderungen ersetzt werden kann.

Auf der Basis von Kapitel II wollen wir nun stochastische Ungleichungen für den U-Prozeß herleiten, die uns in die Lage versetzen werden, L_b -Aussagen bzw. Glivenko-Cantelli-Resultate zu beweisen. Dabei verwenden wir nicht die bei der Untersuchung von U-Statistiken übliche Projektionsmethode (die auch bei NOLAN-POLLARD(1985) benutzt wird), sondern die folgende zentrale Darstellung des U-Prozesses als "Arithmetisches Mittel" von empirischen C -Prozessen (was in dieser Form auf HOEFFDING(1963) zurückgeht; vgl. hierzu auch SILVERMAN(1983)):

3.1.3 Definition. Sei $n \in \mathbb{N}$, $[\cdot]$ bezeichne die Gaußklammer.

$S_n := \{\text{Permutationen von } (1, \dots, i(n))\}.$

Für jedes $\alpha \in S_n$ und $D \in \bigotimes_{\lambda} \mathcal{D}$ sei dann:

$$\mu_n^\alpha(D) := \left[\frac{i(n)}{m} \right]^{-1} \sum_{j=0}^{\left[\frac{i(n)}{m} \right] - 1} 1_D(\xi_{n\alpha(mj+1)}, \dots, \xi_{n\alpha(mj+m)}),$$

$$v_n^\alpha(D) := \mathbb{E}(\mu_n^\alpha(D)),$$

$$\beta_n^\alpha(D) := \left[\frac{i(n)}{m} \right]^{1/2} (\mu_n^\alpha(D) - v_n^\alpha(D)),$$

$$1_n := i(n)^{1/2} \left[\frac{i(n)}{m} \right]^{-1/2} \in [m^{1/2}, m].$$

3.1.4 Lemma. Für alle $n \in \mathbb{N}$ gilt:

(i) Für jedes $\alpha \in S_n$ ist $(\beta_n^\alpha(C))_{C \in \mathcal{C}}$ ein empirischer C-Prozeß zu $\left[\frac{i(n)}{m} \right]$ unabhängigen Zufallselementen in $(\mathbb{Z}^m, \bigotimes_{\lambda} \mathcal{D})$.

$$(ii) H_n = \frac{1}{i(n)!} \sum_{\alpha \in S_n} \mu_n^\alpha.$$

$$(iii) \bar{H}_n = \frac{1}{i(n)!} \sum_{\alpha \in S_n} v_n^\alpha.$$

$$(iv) U_n = 1_n \cdot \frac{1}{i(n)!} \sum_{\alpha \in S_n} \beta_n^\alpha.$$

Beweis. (i) ist unmittelbar klar. Wenn wir (ii) zeigen können, folgt sofort durch Bildung der Erwartungswerte (iii) und somit auch (iv). Sei also $n \in \mathbb{N}$ sowie $D \in \bigotimes_{\lambda} \mathcal{D}$ fest gewählt. Dann ist:

$$\frac{1}{i(n)!} \sum_{\alpha \in S_n} \mu_n^\alpha(D) = \frac{1}{i(n)!} \sum_{\alpha \in S_n} \left[\frac{i(n)}{m} \right]^{-1} \sum_{j=0}^{\left[\frac{i(n)}{m} \right] - 1} 1_D(\xi_{n\alpha(mj+1)}, \dots, \xi_{n\alpha(mj+m)})$$

$$\begin{aligned}
 &= \frac{1}{i(n)!} \left[\frac{i(n)}{m} \right]^{-1} \sum_{j=0}^{\left[\frac{i(n)}{m} \right] - 1} \sum_{\alpha \in S_n} 1_D(\xi_{n\alpha(mj+1)}, \dots, \xi_{n\alpha(mj+m)}) \\
 &= \frac{1}{i(n)!} \left[\frac{i(n)}{m} \right]^{-1} \sum_{j=0}^{\left[\frac{i(n)}{m} \right] - 1} \sum_{(i_1, \dots, i_m) \in E_n} 1_D(\xi_{ni_1}, \dots, \xi_{ni_m}) \\
 &\quad \alpha(mj+k) = i_k \\
 &\quad \text{für } k=1, \dots, m \\
 &= \frac{1}{i(n)!} \left[\frac{i(n)}{m} \right]^{-1} \sum_{j=0}^{\left[\frac{i(n)}{m} \right] - 1} \sum_{(i_1, \dots, i_m) \in E_n} [i(n)-m]! 1_D(\xi_{ni_1}, \dots, \xi_{ni_m}) \\
 &= H_n(D). \quad \square
 \end{aligned}$$

3.1.5 Bemerkung. Durch die Darstellung in 3.1.4 sind wir in der Lage, auch im nicht identisch verteilten Fall (im Gegensatz zu NOLAN-POLLARD(1985)) zu Resultaten zu kommen.

Überdies kann auf die Voraussetzung

C ist eine "symmetrische" Vapnik-Chervonenkis-Klasse, d.h. für $C \in C$ und jede Permutation $(i_1, \dots, i_m)$ von $(1, \dots, m)$ gilt: mit $(z_1, \dots, z_m)$ ist stets auch $(z_{i_1}, \dots, z_{i_m})$ ein Element aus C

verzichtet werden (vgl. hierzu NOLAN-POLLARD(1985), wo symmetrische Funktionenklassen untersucht werden).

3.2 Stochastische Ungleichungen

Wie schon in 2.1.5(ii) angedeutet, wollen wir nun (unter Verwendung von 2.1.4(ii) bzw. 2.1.7) entsprechende Ungleichungen für den U-Prozeß herleiten. Dabei wird Lemma 3.1.4

eine ganz entscheidende Rolle spielen. Es ist natürlich klar, daß sich die "Tail"-Wahrscheinlichkeiten von 2.1.4(i) bzw. 2.1.6 i.a. nicht auf das "Arithmetische Mittel" U_n übertragen lassen, da nicht gilt:

$$\mathbb{P}(g(\frac{U_n}{1_n}) \geq \epsilon) \leq \frac{1}{1(n)!} \sum_{\alpha \in S_n} \mathbb{P}(g(\beta_n^\alpha) \geq \epsilon) \text{ für eine meßbare Abbildung } g: \ell^\infty(C) \rightarrow \mathbb{R}, \text{ selbst dann nicht, wenn } g \text{ konvex ist.}$$

Hingegen gilt für jede konvexe und meßbare Abbildung $g: \ell^\infty(C) \rightarrow \mathbb{R}$ natürlich:

$$\mathbb{E}(g(\frac{U_n}{1_n})) \leq \frac{1}{1(n)!} \sum_{\alpha \in S_n} \mathbb{E}(g(\beta_n^\alpha)),$$

so daß sich die Ergebnisse von 2.1.4(ii) und 2.1.7 auf U_n übertragen lassen. Zunächst interessieren wir uns für die zu 2.1.7 analoge Aussage:

3.2.1 Theorem. Für alle $0 < s < 2$ existiert eine nur von s und C abhängige Konstante K mit der Eigenschaft:

$$\mathbb{E}(\exp((2-s) \|\frac{i(n)}{m}\| R_n^2)) \leq K \text{ für alle } n \in \mathbb{N}.$$

Dabei bezeichne $\|\cdot\|$ wieder die Gaußklammer und für $n \in \mathbb{N}$ sei $R_n := \sup\{|H_n(C) - \bar{H}_n(C)| : C \in C\}$.

Beweis. Gemäß 2.1.7 und 3.1.4(i) existiert zu $0 < s < 2$ ein K (nur abhängig von s und C) mit:

$$(*) \quad \mathbb{E}(\exp((2-s) \|\frac{i(n)}{m}\| D_{n\alpha}^2)) \leq K \text{ für alle } n \in \mathbb{N} \text{ und } \alpha \in S_n,$$

$$\text{wobei } D_{n\alpha} := \sup\{|\mu_n^\alpha(C) - \nu_n^\alpha(C)| : C \in C\}.$$

Wegen 3.1.4 folgt dann sofort:

$$R_n \leq \frac{1}{i(n)!} \sum_{\alpha \in S_n} D_{n\alpha},$$

d.h. mit (*) und wegen der Monotonie sowie der Konvexität von $x \mapsto \exp(x^2)$ ergibt sich:

$$\mathbb{E}(\exp((2-s) \left\| \frac{i(n)}{m} \right\| R_n^2)) \leq \frac{1}{i(n)!} \sum_{\alpha \in S_n} \mathbb{E}(\exp((2-s) \left\| \frac{i(n)}{m} \right\| D_{n\alpha}^2))$$

$\leq K$ für alle $n \in \mathbb{N}$.

□

Mittels Markoffungleichung folgt aus 3.2.1 unmittelbar:

3.2.2 Korollar. Für alle $0 < s < 2$ existiert eine Konstante K , die nur von s und C abhängt, mit der Eigenschaft:

$$\mathbb{P}(R_n \geq \epsilon) \leq K \cdot \exp(-(2-s) \left\| \frac{i(n)}{m} \right\| \epsilon^2) \quad \text{für alle } n \in \mathbb{N}, \epsilon > 0.$$

Dabei ist R_n definiert wie in 3.2.1.

3.2.3 Bemerkung. Ein Vergleich mit der Ungleichung von KIEFER(1961) (vgl. auch GÄNSSLER-STUTE(1979), S.225) zeigt, daß 3.2.2 bis auf die Konstante K die bestmögliche Abschätzung dieser Art darstellt. 3.2.2 enthält auch die exponentielle Schranke aus HELMERS-JANSSEN-SERFLING(1985). In der Situation unserer Bemerkung 3.1.2(ii) wird dort im Theorem 2.1 gezeigt: Zu jedem $0 < s < 2$ existiert ein K in Abhängigkeit von s mit:

$$\mathbb{P}(R_n > \epsilon) \leq K \exp(-(2-s) \left\| \frac{n}{m} \right\| \epsilon^2) \quad \text{für alle } \epsilon > 0 \text{ und } n \in \mathbb{N}.$$

Damit ist aber unmittelbar klar, daß dieses Resultat eine Folgerung ist aus 3.2.2, da die hier betrachtete Vapnik-Chervonenkis-Klasse C wegen ihrer Totalordnung die VC-Zahl

$V(C)=2$ besitzt (K aus 3.2.2 hängt lediglich von $V(C)$ ab; man beachte 2.1.6). Als Konsequenz aus 3.2.2 kann in dieser Situation insbesondere auf die Voraussetzung der identischen Verteiltheit verzichtet werden.

Wir wollen nun für U_n auch ein Analogon zu 2.1.4 herleiten. Dazu betrachten wir für alle $n \in \mathbb{N}$, $\delta > 0$ und $\alpha \in S_n$:

$$\tilde{\omega}_{U_n}(\delta) := \sup\{ |U_n(C) - U_n(D)| : C, D \in \mathcal{C}; \bar{H}_n(C \Delta D) \leq \delta \},$$

$$\omega_n^\alpha(\delta) := \sup\{ |\beta_n^\alpha(C) - \beta_n^\alpha(D)| : C, D \in \mathcal{C}; \nu_n^\alpha(C \Delta D) \leq \delta \},$$

$$M_n^\alpha(\delta) := 2 \sup\{ |\beta_n^\alpha(C)| : C \in \mathcal{C} \}.$$

Ganz entscheidend für das weitere Vorgehen ist nun das folgende Lemma:

3.2.4 Lemma. Für alle $n \in \mathbb{N}$ und $\delta_1, \delta_2 > 0$ gilt:

$$\tilde{\omega}_{U_n}(\delta_1) \leq 1_n \cdot \left[\frac{1}{i(n)!} \sum_{\alpha \in S_n} \omega_n^\alpha(\delta_2) + \left[\frac{1}{i(n)!} \sum_{\alpha \in S_n} (M_n^\alpha)^2 \cdot \delta_1 \cdot \delta_2^{-1} \right]^{1/2} \right].$$

Beweis. Sei $n \in \mathbb{N}$ und $\delta_1, \delta_2 > 0$ fest gewählt. Dann gilt:

$$\begin{aligned} \tilde{\omega}_{U_n}(\delta_1) &\stackrel{3.1.4}{\leq} 1_n \cdot \sup\{ \frac{1}{i(n)!} \sum_{\alpha \in S_n} |\beta_n^\alpha(C) - \beta_n^\alpha(D)| : \bar{H}_n(C \Delta D) \leq \delta_1; C, D \in \mathcal{C} \} \\ &\leq 1_n \cdot \sup\{ \frac{1}{i(n)!} \sum_{\alpha \in S_n} \{ \omega_n^\alpha(\delta_2) + M_n^\alpha \cdot 1(\nu_n^\alpha(C \Delta D) > \delta_2) \} : \bar{H}_n(C \Delta D) \leq \delta_1; C, D \in \mathcal{C} \} \\ &\leq 1_n \left[\frac{1}{i(n)!} \sum_{\alpha \in S_n} \omega_n^\alpha(\delta_2) + A_n(\delta_1, \delta_2) \right], \end{aligned}$$

wobei $A_n(\delta_1, \delta_2) := \sup\{ \frac{1}{i(n)!} \sum_{\alpha \in S_n} M_n^\alpha \cdot 1(\nu_n^\alpha(C \Delta D) > \delta_2) : \bar{H}_n(C \Delta D) \leq \delta_1; C, D \in \mathcal{C} \}.$

Hieraus folgt nun sofort mittels "Cauchy-Schwarz-Ungleichung" die Behauptung, denn es gilt:

$$A_n(\delta_1, \delta_2) \leq \sup \left\{ \left[\frac{1}{i(n)!} \sum_{\alpha \in S_n} (M_n^\alpha)^2 \right]^{1/2} \cdot \left[\frac{1}{i(n)!} \sum_{\alpha \in S_n} 1 (v_n^\alpha(CAD) > \delta_2) \right]^{1/2} \right\}.$$

$$\bar{H}_n(CAD) \leq \delta_1; \quad C, D \in \mathcal{C}$$

$$\leq \left[\frac{1}{i(n)!} \sum_{\alpha \in S_n} (M_n^\alpha)^2 \cdot \delta_1 \cdot \delta_2^{-1} \right]^{1/2}. \quad \square$$

3.2.5 Theorem. Für alle $0 < s < \frac{1}{4}$ existieren Konstanten K und R (die nur von s und C abhängen) so, daß für alle $0 < \delta \leq 1$ und $n \in \mathbb{N}$ mit $i(n) \geq \frac{3}{2} m \delta^{-1}$ gilt:

$$\delta^{1+2V(C^-)} \mathbb{E}(\exp(R\delta^{s-1/4} \tilde{\omega}_{U_n}(\delta))) \leq K.$$

Beweis. Sei $0 < s < \frac{1}{4}$ fest gewählt. Dann definieren wir $x := \frac{1}{4} - s$. Wegen $\frac{x}{1-2x} \in (0, \frac{1}{2})$ existieren gemäß 2.1.4(ii) und 2.1.7 Konstanten R und K (die nur von s und C abhängen) mit der Eigenschaft:

$$\delta^{1+2V(C^-)} \mathbb{E}(\exp(R\delta^{-\frac{x}{1-2x}} \omega_n^\alpha(\delta))) \leq K \quad \text{sowie}$$

$$\mathbb{E}(\exp(R(M_n^\alpha)^2)) \leq K \quad \text{für alle } \alpha \in S_n, \quad 0 < \delta \leq 1 \quad \text{und} \quad \left\lfloor \frac{i(n)}{m} \right\rfloor \geq (2\delta)^{-1}.$$

Aus 3.2.4 erhalten wir außerdem für alle $0 < \delta \leq 1$ und $n \in \mathbb{N}$ (man setze $\delta_1 = \delta$ sowie $\delta_2 = \delta^{1-2x}$):

$$\tilde{\omega}_{U_n}(\delta) \leq 1_n \cdot \left[\frac{1}{i(n)!} \sum_{\alpha \in S_n} \omega_n^\alpha(\delta^{1-2x}) + \left[\frac{1}{i(n)!} \sum_{\alpha \in S_n} (M_n^\alpha)^2 \cdot \delta^{2x} \right]^{1/2} \right].$$

Insgesamt liefert dies nun für alle $0 < \delta \leq 1$ und $n \in \mathbb{N}$ mit

$i(n) \geq \frac{3}{2} m \delta^{-1}$ (also $\lfloor \frac{i(n)}{m} \rfloor \geq \frac{i(n)}{m} - 1 \geq \frac{3}{2} \delta^{-1} - 1 \geq (2\delta^{1-2x})^{-1}$) wegen der Monotonie und Konvexität von $x \mapsto \exp(x)$:

$$\begin{aligned} 2 \mathbb{E}(\exp(\frac{R}{2m} \delta^{-x} \tilde{\omega}_{U_n}(\delta))) &\leq \frac{1}{i(n)!} \sum_{\alpha \in S_n} \mathbb{E}(\exp(R(\delta^{1-2x})^{-\frac{x}{1-2x}} \omega_n^\alpha(\delta^{1-2x}))) \\ &\quad + \frac{1}{i(n)!} \sum_{\alpha \in S_n} \mathbb{E}(\exp(R[(M_n^\alpha)^2 + 1])) \\ &\leq K(\delta^{1-2x})^{-1-2V(C^-)} + \exp(R) \cdot K \\ &\leq 2\exp(R) K \delta^{-1-2V(C^-)} . \quad \square \end{aligned}$$

Mittels Markoffungleichung erhalten wir aus 3.2.5 ($s = \frac{1}{20}$):

3.2.6 Korollar. Es existieren Konstanten K, K' und R (nur abhängig von C) so, daß für alle $0 < \delta \leq 1$, $0 < \varepsilon$ sowie $n \in \mathbb{N}$ mit $i(n) \geq \frac{3}{2} m \delta^{-1}$ gilt:

$$\begin{aligned} \mathbb{P}(\tilde{\omega}_{U_n}(\delta) \geq \varepsilon) &\leq K \delta^{-1-2V(C^-)} \exp(-R\varepsilon \delta^{-1/5}) \\ &\leq K' \varepsilon^{-5-10V(C^-)} \exp(-\frac{R}{2} \varepsilon \delta^{-1/5}) . \end{aligned}$$

Sehr wesentlich für den weiteren Verlauf wird das Lemma 3.2.7 sein. Dazu führen wir den Hilfsprozeß L_n ein, der für jedes $n \in \mathbb{N}$ und $D \in \bigotimes_{\lambda} \mathcal{D}$ wie folgt definiert ist:

$$L_n(D) := \sum_{j=1}^m \left[\sum_{i=1}^{j-1} X_i v_n \times \beta_n \times \sum_{j=1}^m X_j v_n \right] (D) , \text{ wobei für } B \in \mathcal{D}$$

$$\mu_n(B) := \frac{1}{i(n)} \sum_{i=1}^{i(n)} 1_B(\xi_{ni}) ,$$

$$v_n(B) := \mathbb{E}(\mu_n(B)) \text{ und } \beta_n(B) := i(n)^{1/2} [\mu_n(B) - v_n(B)].$$

Dabei bezeichne $\times$ die Multiplikation der Maße und wir stellen fest, daß β_n nichts anderes ist als der empirische Prozeß zu $\xi_{n1}, \dots, \xi_{ni(n)}$. Mit diesen Bezeichnungen folgt nun:

3.2.7 Lemma. Für alle $k \in \mathbb{N}$ existiert eine positive Konstante $C(k, m)$ (d.h. nur abhängig von k und m) mit der Eigenschaft:

$$\mathbb{E}(|U_n(D) - L_n(D)|^{2k}) \leq C(k, m) \cdot i(n)^{-k} \text{ für alle } n \in \mathbb{N}, D \in \bigotimes_1^m \mathcal{D}.$$

Beweis. Wir betrachten lediglich die Situation $m \geq 2$ ($m=1$ ist trivial, da dann $U_n = L_n$). Zunächst definieren wir für alle $n \in \mathbb{N}$ und $D \in \bigotimes_1^m \mathcal{D}$:

$$\bar{U}_n(D) := \sum_{j=1}^m \left[\frac{1}{X} v_n \times \beta_n \times \frac{1}{j+1} \sum_{j+1}^m \mu_n \right] (D) = i(n)^{1/2} \left[\frac{1}{X} \mu_n - \frac{1}{X} v_n \right] (D).$$

Dann gilt für alle $n \in \mathbb{N}$:

$$\begin{aligned} |U_n - L_n| &\leq |U_n - \bar{U}_n| + |\bar{U}_n - L_n| \\ &\leq 2 \left[1 - \prod_{i=0}^{m-1} \frac{i(n)-i}{i(n)} \right] i(n)^{1/2} + |\bar{U}_n - L_n| \\ &\leq 2 m^2 i(n)^{-1/2} + |\bar{U}_n - L_n|. \end{aligned}$$

Da außerdem gilt:

$$|\bar{U}_n - L_n| \leq \sum_{j=1}^{m-1} \left| \frac{1}{X} v_n \times \beta_n \times \left[\frac{1}{j+1} \sum_{j+1}^m \mu_n - \frac{1}{j+1} \sum_{j+1}^m v_n \right] \right|,$$

folgt die Behauptung von 3.2.7, falls für $k \in \mathbb{N}$ gilt:

$$(*) \quad \mathbb{E} \left(\left| \beta_n \times \left(\frac{1}{X} \mu_n - \frac{1}{X} v_n \right) \right| (D) \right)^{2k} \leq C(k, r) i(n)^{-k} \text{ für alle } r \in \mathbb{N}, D \in \bigotimes_1^{r+1} \mathcal{D} \text{ sowie } n \in \mathbb{N}.$$

Beweis zu (*): Erfolgt durch Induktion nach r .

$r=1$: Sei $n \in \mathbb{N}$ sowie $D \in \bigotimes_1^{2k} \mathcal{D}$ beliebig gewählt. Mit $E_k := \{1, \dots, i(n)\}^{4k}$ und $v_{ni} := L(\xi_{ni})$ für $i \leq i(n)$ folgt dann:

$$\begin{aligned} \mathbb{E}([\beta_n \times (\mu_n - v_n)](D)^{2k}) &= \\ i(n)^{-3k} \sum_{(i_1^1, i_1^2, \dots, i_1^{2k}, i_2^{2k}) \in E_k} \mathbb{E}(\prod_{j=1}^{2k} [(1_{\xi_{ni_1^j}}^{-v_{ni_1^j}}) \times (1_{\xi_{ni_2^j}}^{-v_{ni_2^j}})](D)) \\ &\leq i(n)^{-3k} \sum_{(i_1, \dots, i_{4k}) \in E_k \text{ mit } |\{i_1, \dots, i_{4k}\}| \leq 2k} 2^{2k} \\ &\leq i(n)^{-3k} \sum_{j=1}^{2k} \binom{4k}{j} i(n)^j j^{4k-j} 2^{2k} \leq C(k) i(n)^{-k}. \end{aligned}$$

$r \mapsto r+1$:

Wegen $\beta_n \times [\begin{smallmatrix} r+1 & r+1 \\ X & X \\ \mu_n & v_n \end{smallmatrix}] = \beta_n \times [\begin{smallmatrix} r & r \\ X & X \\ \mu_n & -X v_n \end{smallmatrix}] \times \mu_n + \beta_n \times [\begin{smallmatrix} r \\ X v_n \end{smallmatrix}] \times [\begin{smallmatrix} r \\ \mu_n - v_n \end{smallmatrix}]$ folgt für alle $n \in \mathbb{N}$ und $D \in \bigotimes_1^{r+1} \mathcal{D}$:

$$\begin{aligned} \mathbb{E}([\beta_n \times [\begin{smallmatrix} r+1 & r+1 \\ X & X \\ \mu_n & -X v_n \end{smallmatrix}]](D)^{2k}) &\leq \\ 2^{2k} [i(n)^{-1} \sum_{i=1}^{i(n)} \mathbb{E}([\beta_n \times [\begin{smallmatrix} r & r \\ X & X \\ \mu_n & -X v_n \end{smallmatrix}]] \times 1_{\xi_{ni}})](D)^{2k}) &+ \\ \mathbb{E}([\beta_n \times [\begin{smallmatrix} r \\ X v_n \end{smallmatrix}] \times (\mu_n - v_n)](D)^{2k}) &]. \end{aligned}$$

Den ersten Summanden können wir nun geeignet mit der Induktionsannahme abschätzen, während für den zweiten die Situation $r=1$ zutrifft. Damit ist aber die Induktionsbehauptung gezeigt (und auch 3.2.7). $\square$

3.2.8 Lemma. Für alle $k \in \mathbb{N}$ existiert eine Konstante

$C(k, m)$ so, daß für alle $n \in \mathbb{N}$ und $C, D \in \bigotimes_1^m \mathcal{D}$ gilt:

$$\mathbb{E}([U_n(C) - U_n(D)]^{2k}) \leq C(k, m) [\mathbb{E}([L_n(C) - L_n(D)]^2)^k + i(n)^{-k}].$$

Beweis. Sei $k \in \mathbb{N}$ fest vorgegeben. Für beliebiges $n \in \mathbb{N}$

sowie $C, D \in \bigotimes_1^m \mathcal{D}$ betrachten wir:

$$r := \mathbb{E}([L_n(C) - L_n(D)]^2) \quad \text{und} \quad t := \left(\frac{3ri(n)^{1/2}}{2m}\right)^{2k}.$$

Mithilfe der Bernsteinungleichung (GÄNSSLER(1983), S.9)

folgt nun:

$$\begin{aligned} \mathbb{E}([L_n(C) - L_n(D)]^{2k}) &= \int_0^\infty \mathbb{P}(|L_n(C) - L_n(D)| \geq \varepsilon^{\frac{1}{2k}}) d\varepsilon \\ &\leq \int_0^\infty 2 \exp\left(-\frac{3\varepsilon^{1/k}}{6r + 4\varepsilon^{\frac{1}{(2k)} m \cdot i(n)^{-1/2}}}\right) d\varepsilon \\ &\leq 2 \left[\int_0^t \exp\left(-\frac{\varepsilon^{1/k}}{4r}\right) d\varepsilon + \int_t^\infty \exp\left(-\frac{3\varepsilon^{1/(2k)} i(n)^{1/2}}{8m}\right) d\varepsilon \right] \\ &\leq 2 \left[r^k \int_0^t \exp\left(-\frac{1}{4} u^{\frac{1}{k}}\right) du + i(n)^{-k} \int_{ti(n)^k}^\infty \exp\left(-\frac{3}{8m} u^{\frac{1}{2k}}\right) du \right] \\ &\leq C(k, m) [r^k + i(n)^{-k}]. \end{aligned}$$

Mithilfe von 3.2.7 folgt sofort die Behauptung. □

3.2.9 Bemerkung. Für die Herleitung der L_b -Aussagen in 3.3

wird sehr wesentlich sein, daß gilt:

$$\mathbb{E}([L_n(C) - L_n(D)]^2) \leq 2m^2 \left[\sum_1^m v_n \right] (C \Delta D) \leq 2m^2 [\bar{H}_n(C \Delta D) + 2m^2 i(n)^{-1}]$$

für alle $n \in \mathbb{N}$ und $C, D \in \bigotimes_1^m \mathcal{D}$.

3.3 L_b -Konvergenz und Glivenko-Cantelli-Theorem

Ähnlich wie in Kapitel II stellt 3.2.6 den Zugang zur relativen L_b -Folgenkompaktheit von $(U_n)_{n \in \mathbb{N}}$ dar, vorausgesetzt wir stellen eine zur dortigen Bedingung (A) analoge Forderung. Dabei wird uns 3.2.8 helfen, diese Bedingung bzgl. der Varianzstruktur des Prozesses zu formulieren (mit 3.2.6 könnten wir dies zunächst nur bzgl. $(\bar{H}_n)_{n \in \mathbb{N}}$ tun). Es möge also in unseren folgenden L_b -Aussagen stets gelten:

(C) Es existiert eine Halbmetrik s auf C mit:

$$(i) \lim_{n \rightarrow \infty} \sup \sup \{ \mathbb{E}([L_n(C) - L_n(D)]^2) : s(C, D) \leq \delta; C, D \in C \} \rightarrow 0$$

für $\delta \downarrow 0$.

(ii) (C, s) ist totalbeschränkt.

Falls dann (C) erfüllt ist, benutzen wir in den L_b -Aussagen immer das Modell aus Kapitel I mit:

$$X=C; s \text{ wie in (C); } A=\sigma(\pi_C(S): S \in C) \text{ mit } S=\ell^\infty(C).$$

Dann können wir $(U_n)_{n \in \mathbb{N}}$ auffassen als eine Folge von F, A -meßbaren Zufallselementen in S und es ist klar, daß wegen der Abzählbarkeit von C die Meßbarkeitsforderungen aus Kapitel I erfüllt sind, d.h. die dortigen Ergebnisse sind anwendbar.

Bevor wir uns der L_b -Konvergenz zuwenden, wollen wir noch das folgende Glivenko-Cantelli-Resultat beweisen:

3.3.1 Theorem. Sei $(\gamma_n)_{n \in \mathbb{N}} \subset (1, \infty)$ mit $\sum_{n=1}^{\infty} \gamma_n^{-\varepsilon} < \infty$ für alle $\varepsilon > 0$. Dann gilt:

$$[\log \gamma_n]^{-1/2} \sup \{ |U_n(C)| : C \in C \} \rightarrow 0 \quad \mathbb{P}\text{-f.s.}$$

Beweis. Für alle $n \in \mathbb{N}$ und $\epsilon > 0$ gilt gemäß 3.2.2:

$$\begin{aligned} \mathbb{P}([\log \gamma_n]^{-1/2} \sup\{ |U_n(C)| : C \in \mathcal{C} \} \geq \epsilon) \\ \leq K \exp\left(-\frac{\log \gamma_n}{i(n)} \left[\frac{i(n)}{m} \right] \epsilon^2\right) \\ \leq K \exp\left(-\frac{\epsilon^2}{m^2} \log \gamma_n\right) \leq K \gamma_n^{-(\epsilon/m)^2}. \end{aligned}$$

Aufgrund der vorausgesetzten Reihenbedingung folgt nun mittels Borel-Cantelli sofort die Behauptung. $\square$

3.3.2 Theorem. Sei (C) erfüllt und $i(n) \rightarrow \infty$ für $n \rightarrow \infty$. Dann ist $(U_n)_{n \in \mathbb{N}}$ relativ L_b -folgenkompakt.

Beweis. Für $n \in \mathbb{N}$ und $\delta > 0$ definieren wir:

$$\omega_n^r(\delta) := \sup\{ |U_n(C) - U_n(D)| : r(C, D) \leq \delta; C, D \in \mathcal{C} \} \text{ für } r \in \{s, s_n\},$$

$$\text{wobei } s_n(C, D) := \mathbb{E}([L_n(C) - L_n(D)]^2)^{1/2} \text{ für } C, D \in \mathcal{C}.$$

$$a_n(\delta) := \sup\{s_n(C, D) : s(C, D) \leq \delta; C, D \in \mathcal{C}\}.$$

Dann ergibt sich wie im Beweis von 2.2.4 (man beachte, daß

$$\text{wegen 3.2.9 gilt: } s_n(C, D) \leq 2m[\bar{H}_n(CAD) + 2m^2 i(n)^{-1}]^{1/2})$$

für alle $n \in \mathbb{N}$ und $0 < \delta \leq 1$:

$$\omega_n^{s_n}(\delta) \leq 2\tilde{\omega}_{U_n}(\delta) + R_n(\delta), \text{ wobei } \tilde{\omega}_{U_n} \text{ wie in 3.2.4 und}$$

$$R_n(\delta) := \sup\{ |U_n(C) - U_n(D)| : s_n(C, D) \leq 3m[\delta + 2m^2 i(n)^{-1}]^{1/2} + \delta; C, D \in N_n(\delta) \}$$

$$\text{mit } N_n(\delta) \subset \bigotimes_1^m \mathcal{D}, |N_n(\delta)| \leq K(C) \delta^{-1-V(C)}.$$

Insgesamt folgt nun für alle $\epsilon > 0$ (wobei K eine variable, höchstens von m und C abhängige Konstante bezeichne):

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(\omega_n^s(\delta) \geq \epsilon) \stackrel{3.2.6}{\leq} \lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(R_n(\delta) \geq \frac{\epsilon}{2}) \leq$$

$$\limsup_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} K \delta^{-2-2V(C)} \sup\{\mathbb{P}(|U_n(C) - U_n(D)| \geq \frac{\epsilon}{2}) : s_n(C, D) \leq$$

$$\delta + 3m[\delta + 2m^2 i(n)^{-1}]^{1/2}; C, D \in \bigotimes_1^m \mathcal{D}\} \leq$$

$$\limsup_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} K \delta^{-2-2V(C)} \sup\{(\frac{\epsilon}{2})^{-6-4V(C)} \mathbb{E}([U_n(C) - U_n(D)]^{6+4V(C)}) :$$

$$s_n(C, D) \leq \delta + 4m\delta^{1/2}; C, D \in \bigotimes_1^m \mathcal{D}\} \leq$$

$$\stackrel{3.2.8}{\leq} K \epsilon^{-6-4V(C)} \limsup_{\delta \downarrow 0} \delta^{-2-2V(C)} [\delta^{3+2V(C)} + \delta^{6+4V(C)}] = 0.$$

Da außerdem für alle $n \in \mathbb{N}$ und $0 < \delta \leq 1$ gilt: $\omega_n^s(\delta) \leq \omega_n^{s_n}(a_n(\delta))$,
erhalten wir sofort aus der Gültigkeit von (C)(i):

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(\omega_n^s(\delta) \geq \epsilon) = 0 \quad \text{für alle } \epsilon > 0, \text{ womit also}$$

(iii)(A) aus Theorem 1.3.4 erfüllt ist. (iii)(B) hingegen
ist eine unmittelbare Folgerung aus der Markoffungleichung,
denn mit 3.2.9 erhalten wir für alle $C \in \mathcal{C}$ und $K > 0$:

$$\mathbb{P}(|U_n(C)| \geq K) \leq K^{-2} \mathbb{E}(U_n(C)^2) \leq K(m) K^{-2}, \text{ wobei } K(m) \text{ eine}$$

nur von m abhängige Konstante darstellt. □

3.3.3 Theorem. Sei (C) erfüllt sowie $i(n) \rightarrow \infty$ für $n \rightarrow \infty$.

Dann sind äquivalent:

(i) Es existiert ein zentrierter Gaußprozeß G mit $U_n \xrightarrow{L_b} G$.

(ii) Es gilt (i) mit $\mathbb{E}(G(C)G(D)) = \lim_{n \rightarrow \infty} \mathbb{E}(U_n(C)U_n(D))$ für $C, D \in \mathcal{C}$.

(iii) $(\mathbb{E}(U_n(C)U_n(D)))_{n \in \mathbb{N}}$ ist konvergent für alle $C, D \in \mathcal{C}$.

Beweis.

"(i) $\Rightarrow$ (ii)" Aus $U_n \xrightarrow{L_b} G$ folgt mit 1.3.3 für alle $C, D \in C$:

$$(*) \quad (U_n(C), U_n(D)) \xrightarrow{L} (G(C), G(D)).$$

Da außerdem für alle $n \in \mathbb{N}$ wegen 3.2.8 bzw. 3.2.9 gilt:

$$\mathbb{E}(U_n(C)^2 U_n(D)^2) \leq [\mathbb{E}(U_n(C)^4) \mathbb{E}(U_n(D)^4)]^{1/2} \leq K(m) \text{ mit}$$

einer nur von m abhängigen Konstanten $K(m)$, ist also

$(U_n(C) U_n(D))_{n \in \mathbb{N}}$ gleichgradig integrierbar und wir erhalten sofort aus (*):

$$\mathbb{E}(U_n(C) U_n(D)) \longrightarrow \mathbb{E}(G(C) G(D)) \text{ für } n \rightarrow \infty.$$

"(ii) $\Rightarrow$ (iii)" Klar.

"(iii) $\Rightarrow$ (i)" Für $n \in \mathbb{N}$, $i=1, \dots, i(n)$ und $j=1, \dots, m$ definieren wir:

$$D_{ni}^j := \prod_{l=1}^{j-1} v_n \times (1_{\xi_{ni}}^{-L(\xi_{ni})}) \times \prod_{j+1}^m v_n.$$

Seien nun $C_1, \dots, C_k \in C$, $k \in \mathbb{N}$ fest vorgegeben, und sei weiter $(N_1, \dots, N_k)$ ein normalverteilter Zufallsvektor mit

$$\mathbb{E}(N_i N_j) = \lim_{n \rightarrow \infty} \mathbb{E}(U_n(C_i) U_n(C_j)) \text{ für } i, j=1, \dots, k \text{ (der gemäß}$$

GÄNSSLER-STUTE(1977), 1.19.12, existiert). Wir zeigen nun,

daß für eine beliebige Wahl $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ gilt:

$$(**) \quad \sum_{i=1}^k \lambda_i L_n(C_i) \xrightarrow{L} \sum_{i=1}^k \lambda_i N_i, \text{ wobei } L_n \text{ der Hilfsprozeß}$$

aus 3.2.7 ist. Mit dem Cramér-Wold-Device liefert dies

sofort $(L_n(C_1), \dots, L_n(C_k)) \xrightarrow{L} (N_1, \dots, N_k)$ und daher auf-

grund von 3.2.7 $(U_n(C_1), \dots, U_n(C_k)) \xrightarrow{L} (N_1, \dots, N_k)$.

Damit ist aber die Konvergenz der endlichdimensionalen Randverteilungen des C-Prozesses $(U_n)_{n \in \mathbb{N}}$ gegen mehrdimensionale, zentrierte Normalverteilungen nachgewiesen und es folgt gemäß 1.4.1 die Behauptung. Wir müssen also lediglich zeigen, daß (**) erfüllt ist:

Seien $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ beliebig gewählt. Dann ist für jedes $n \in \mathbb{N}$

$$\sum_{l=1}^k \lambda_l L_n(C_l) = \sum_{i=1}^{i(n)} Y_{ni}^2, \text{ wobei für } i=1, \dots, i(n)$$

$$Y_{ni} := \sum_{l=1}^k \sum_{j=1}^m D_{ni}^j(C_l) \lambda_l^{i(n)-1/2}.$$

Es gilt nun:

(1) $(Y_{ni})_{i \leq i(n), n \in \mathbb{N}}$ ist ein Dreiecksschema von zeilenweise unabhängigen, zentrierten Zufallsvariablen.

$$(2) \sum_{i=1}^{i(n)} \mathbb{E}(Y_{ni}^2) = \mathbb{E}(\left[\sum_{i=1}^{i(n)} Y_{ni} \right]^2) \stackrel{3.2.7}{=} 0$$

$$\sum_{l=1}^k \lambda_l \mathbb{E}(U_n(C_l) U_n(C_l)) + o(i(n)^{-1/2}) \xrightarrow{n \rightarrow \infty} \mathbb{E}(\left(\sum_{i=1}^k \lambda_i N_i \right)^2).$$

$$(3) \sum_{i=1}^{i(n)} \mathbb{E}(|Y_{ni}|^{2+\epsilon}) \leq i(n)^{-\epsilon/2} \left[\sum_{l=1}^k |\lambda_l| \right]^{2+\epsilon} m^{2+\epsilon} \xrightarrow{n \rightarrow \infty} 0 \text{ für } \epsilon > 0.$$

Aus (1), (2) und (3) folgt mit GÄNSSLER-STUTE(1977), 9.2.9, die Gültigkeit von (**). □

3.3.4 Bemerkung.

(i) Aus 3.3.3 ersieht man sofort, daß mögliche "Häufungspunkte von $(U_n)_{n \in \mathbb{N}}$ stets zentrierte Gaußprozesse sind.

(ii) Analog zum Vorgehen in Kapitel II könnte man natürlich diese Ergebnisse übertragen auf einen "gebootstrappten" U-Prozeß.

(iii) Wegen 3.2.7 gilt für alle $C, D \in C$:

$$| \mathbb{E}(U_n(C)U_n(D)) - \mathbb{E}(L_n(C)L_n(D)) | = O(i(n)^{-1/2}).$$

Unter der Voraussetzung: $i(n) \rightarrow \infty$ für $n \rightarrow \infty$, führt also die Untersuchung der Kovarianzstruktur von $(U_n)_{n \in \mathbb{N}}$ auf die Untersuchung derjenigen von $(L_n)_{n \in \mathbb{N}}$, wobei natürlich gilt:

$$\begin{aligned} \mathbb{E}(L_n(C)L_n(D)) &= \sum_{j=1}^m \sum_{k=1}^m \int v_n(C(x,j) \cap D(y,k)) \left[\prod_{l=1}^{m-1} v_n \right](dx) \left[\prod_{l=1}^{m-1} v_n \right](dy) \\ &\quad + i(n)^{-1} \sum_{i=1}^{i(n)} \left[\prod_{l=1}^{j-1} v_n \times v_{ni} \times \prod_{l=j+1}^m v_n \right](C) \left[\prod_{l=1}^{k-1} v_n \times v_{ni} \times \prod_{l=k+1}^m v_n \right](D), \end{aligned}$$

wobei $v_{ni} := L(\xi_{ni})$ und für $x = (x_1, \dots, x_{m-1}) \in \mathbb{Z}^{m-1}$ sowie

$$j=1, \dots, m \quad C(x, j) := \{z \in \mathbb{Z} : (x_1, \dots, x_{j-1}, z, x_{j+1}, \dots, x_{m-1}) \in C\}$$

(analog für $D(y, k)$).

KAPITEL IV: DER GEWICHTETE U-PROZESS AUF DER REELLEN ACHSE

4.1 Grundbegriffe und Definitionen

Sei im folgenden wiederum $(\xi_{ni})_{i \leq i(n), n \in \mathbb{N}}$ ein Dreiecksschema von zeilenweise unabhängigen Zufallselementen in einem meßbaren Raum $(Z, \mathcal{D})$, definiert über einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ und $\mathcal{F}, \mathcal{D}$ -meßbar. Sei weiter für das gesamte Kapitel IV m eine feste natürliche Zahl (wobei o.E. $i(n) \geq m$ für $n \in \mathbb{N}$) sowie $h: Z^m \rightarrow \mathbb{R}$ eine (nicht notwendig symmetrische) Kernfunktion, die $\bigotimes_{i=1}^m \mathcal{D}, \mathcal{B}^*$ -meßbar ist (da bei bezeichne $\mathcal{B}^*$ die Borel- σ -Algebra auf $\mathbb{R}$).

Im weiteren Verlauf sei dann stets $q: [0, 1] \rightarrow \mathbb{R}$ eine Gewichtsfunktion mit den Eigenschaften:

$$\begin{aligned} q(0) &= 0; \quad q \text{ ist strikt wachsend auf } [0, \frac{1}{2}] \text{ und sym-} \\ &\text{metrisch um } \frac{1}{2}; \quad q \text{ ist stetig; } \int_0^1 q^{-1}(x) dx < \infty. \end{aligned}$$

4.1.1 Definition. Sei $n \in \mathbb{N}$ sowie $t \in \mathbb{R}$ fest gewählt.

$$E_n := \{(i_1, \dots, i_m) \in \{1, \dots, i(n)\}^m : i_j \neq i_k \text{ für } j \neq k \text{ mit } j, k = 1, \dots, m\},$$

$$H_n(t) := |E_n|^{-1} \sum_{(i_1, \dots, i_m) \in E_n} 1(h(\xi_{ni_1}, \dots, \xi_{ni_m}) \leq t),$$

$$\bar{H}_n(t) := \mathbb{E}(H_n(t)),$$

$$U_n(t) := i(n)^{1/2} [H_n(t) - \bar{H}_n(t)].$$

4.1.2 Bemerkung. 4.1.1 ist analog zu 3.1.1, wenn wir als

Vapnik-Chervonenkis-Klasse wählen: $C := \{h^{-1}((-\infty, t]): t \in \mathbb{R}\}$.

Wir wollen nun die relative L_b -Folgenkompaktheit von $(q^{-1}(\bar{H}_n) U_n)_{n \in \mathbb{N}}$ untersuchen (mit der Konvention, daß für $\bar{H}_n(t)=0,1$ $q^{-1}(\bar{H}_n(t))U_n(t):=0$), wobei wir o.E. (da dies mit Wahrscheinlichkeit 1 der Fall ist) annehmen: $\sup\{q^{-1}(\bar{H}_n(t))|U_n(t)|:t \in \mathbb{R}\} < \infty$. Wir sagen, daß q die Bedingung (Q_k) für ein $k \in \mathbb{N}$ erfüllt, falls gilt:

$$\int_0^{1/2} q^{-2k}(x) x^{k-1} dx < \infty.$$

In diesem Fall definieren wir für $0 \leq s \leq \frac{1}{2}$:

$$I(s) := \int_0^s q^{-2k}(x) x^{k-1} dx \quad \text{sowie}$$

$$\bar{I}(s) := \int_0^s [q^{-2k}(x) d(x^k) + x^k d(-q^{-2k}(x))] \leq 2kI(s).$$

Für die Herleitung von L_b -Aussagen werden wir zunächst in 4.2 geeignete stochastische Ungleichungen zur Verfügung stellen. Dabei gehen wir ähnlich vor wie bei SHORACK(1979) und definieren für jedes $n \in \mathbb{N}$ rekursiv $(\bar{H}_n(-\infty):=0=:U_n(-\infty))$: $-\infty =: t_0^n < t_1^n < \dots < t_{m_n}^n < t_{m_n+1}^n := \infty$, wobei für $i=0, \dots, m_n$ $t_{i+1}^n := \inf\{t > t_i^n : \bar{H}_n(t) \geq \bar{H}_n(t_i^n) + i(n)^{-1}\}$ und $\inf(\emptyset) := \infty$.

4.2 Stochastische Ungleichungen

Sei im Verlauf von 4.2 stets die Bedingung (Q_k) für ein festes $k \in \mathbb{N}$ erfüllt. Zunächst gilt dann $(q^{-1}(\bar{H}_n(-\infty))U_n(-\infty):=0)$:

4.2.1 Lemma. Es existiert eine Konstante $C(k,m)$ (also nur abhängig von k und m), so daß für $0 < \delta \leq \frac{1}{2}$, $0 < \epsilon$ und $n \in \mathbb{N}$ gilt:

$$\mathbb{P}(\max\{|q^{-1}(\bar{H}_n(t_i^n))U_n(t_i^n)| : \bar{H}_n(t_i^n) \geq \delta\} \geq \epsilon) \leq C(k,m) \epsilon^{-4k} \bar{I}(\delta)^2.$$

Beweis. Wir machen wieder Gebrauch von 3.2.8 bzw. 3.2.9, wonach gilt: Es existiert eine nur von k und m abhängige Konstante $C(k,m)$ mit der Eigenschaft (*):

$$\mathbb{E}([U_n(t) - U_n(s)]^{4k}) \leq C(k,m) [(\bar{H}_n(t) - \bar{H}_n(s))^{2k} + i(n)^{-2k}]$$

für alle $s, t \in \mathbb{R}$ und $n \in \mathbb{N}$.

Dann gilt für jedes $n \in \mathbb{N}$, $0 < \delta \leq \frac{1}{2}$ sowie $0 \leq i < j \leq m_n$ mit $\bar{H}_n(t_j^n) \leq \delta$ (man beachte, daß $\bar{H}_n(t_j^n) - \bar{H}_n(t_i^n) \geq i(n)^{-1}$):

$$\mathbb{E}([q^{-1}(\bar{H}_n(t_j^n))U_n(t_j^n) - q^{-1}(\bar{H}_n(t_i^n))U_n(t_i^n)]^{4k}) \leq$$

$$2^{4k} [q^{-4k}(\bar{H}_n(t_j^n)) \mathbb{E}([U_n(t_j^n) - U_n(t_i^n)]^{4k}) +$$

$$\mathbb{E}(U_n(t_i^n)^{4k}) [q^{-1}(\bar{H}_n(t_i^n)) - q^{-1}(\bar{H}_n(t_j^n))]^{4k}] \stackrel{(*)}{\leq}$$

$$2^{5k} C(k,m) [q^{-2k}(\bar{H}_n(t_j^n)) [\bar{H}_n(t_j^n)^k - \bar{H}_n(t_i^n)^k]]^2 +$$

$$2^{5k} C(k,m) [\bar{H}_n(t_i^n)^k [q^{-2k}(\bar{H}_n(t_i^n)) - q^{-2k}(\bar{H}_n(t_j^n))]]^2 \leq$$

$$2^{5k} C(k,m) [\bar{I}(\bar{H}_n(t_j^n)) - \bar{I}(\bar{H}_n(t_i^n))]^2 =$$

$$2^{5k} C(k,m) \left[\sum_{r=i+1}^j [\bar{I}(\bar{H}_n(t_r)) - \bar{I}(\bar{H}_n(t_{r-1}))] \right]^2.$$

Die Behauptung ergibt sich nun unmittelbar aus Theorem 12.2 in BILLINGSLEY (1968). □

4.2.2 Theorem. Es existiert eine Konstante $C(k,m) > 0$ so, daß für alle $0 < \delta \leq \frac{1}{2}$, $0 < \epsilon$ und $n \in \mathbb{N}$ mit $i(n) \geq 2$ gilt:

$$\mathbb{P}(\sup\{q^{-1}(\bar{H}_n(t)) | U_n(t) | : \bar{H}_n(t) \leq \delta\} \geq \epsilon) \leq C(k,m) [\epsilon^{-1} I(i(n)^{-1})^{1/(2k)} + \epsilon^{-4k} I(\delta)^2].$$

Beweis. Sei $0 < \delta \leq \frac{1}{2}$, $0 < \epsilon$ und $i(n) \geq 2$ fest gewählt. Mit

$E := \{i \geq 1: \bar{H}_n(t_i^n) \leq \delta\}$ folgt dann:

$$\sup\{q^{-1}(\bar{H}_n(t)) | U_n(t) | : \bar{H}_n(t) \leq \delta\} \leq S_1 + \max\{S_2^i : i \in E\} + S_3,$$

$$\text{wobei } S_1 := \sup\{q^{-1}(\bar{H}_n(t)) | U_n(t) | : t < t_1^n\},$$

$$S_2^i := \sup\{q^{-1}(\bar{H}_n(t_i^n)) | U_n(t) - U_n(t_i^n) | : t_i^n \leq t < t_{i+1}^n\}$$

für $i \in E$,

$$S_3 := \max\{q^{-1}(\bar{H}_n(t_i^n)) | U_n(t_i^n) | : i \in E\}.$$

Wir leiten zunächst eine geeignete Abschätzung für S_1 her.

Es gilt:

$$S_1 \leq i(n)^{\frac{1}{2}} \sum_{(i_1, \dots, i_m) \in E_n} q^{-1}(\bar{H}_n(h(\xi_{ni_1}, \dots, \xi_{ni_m}))) 1_{(h(\xi_{ni_1}, \dots, \xi_{ni_m}) < t_1^n)} \\ + i(n)^{\frac{1}{2}} \sup\{q^{-1}(\bar{H}_n(t)) \cdot \bar{H}_n(t) : t < t_1^n\},$$

also wegen $\bar{H}_n(s) \leq i(n)^{-1}$ für alle $s < t_1^n$:

$$\begin{aligned} \mathbb{E}(S_1) &\leq 2 i(n)^{1/2} \int_0^{i(n)^{-1}} q^{-1}(x) dx \\ &\leq 2 i(n)^{1/2} \int_0^{i(n)^{-1}} [x^{-\frac{k-1}{2k}} 1_{(x \leq \frac{1}{i(n)})}] [(q^{-2k}(x) x^{k-1})^{\frac{1}{2k}} 1_{(x \leq \frac{1}{i(n)})}] dx \\ &\stackrel{\text{Hölder}}{\leq} 2 i(n)^{1/2} [\int_0^{i(n)^{-1}} x^{\frac{k-1}{2k-1}} dx]^{\frac{2k-1}{2k}} I(i(n)^{-1})^{\frac{1}{2k}} \\ &= 2 i(n)^{1/2} [\frac{2k-1}{k} i(n)^{-\frac{k}{2k-1}}]^{\frac{2k-1}{2k}} I(i(n)^{-1})^{\frac{1}{2k}} \\ &= 2 [\frac{2k-1}{k}]^{\frac{2k-1}{2k}} I(i(n)^{-1})^{\frac{1}{2k}}. \end{aligned}$$

Wegen 4.2.1 müssen wir also lediglich noch eine geeignete Abschätzung für $\max\{S_2^i: i \in E\}$ herleiten. Sei dazu für jedes $t \in \mathbb{R}$ $U_n(t-) := \lim_{s \uparrow t} U_n(s)$. Dann folgt für $i \in E$:

$$\begin{aligned} S_2^i &\leq q^{-1}(\bar{H}_n(t_i^n)) |U_n(t_{i+1}^n-) - U_n(t_i^n)| + \\ &\quad 2i(n)^{1/2} q^{-1}(\bar{H}_n(t_i^n)) [\bar{H}_n(t_{i+1}^n-) - \bar{H}_n(t_i^n)] \\ &\leq R_{ni} + 2i(n)^{-1/2} q^{-1}(i(n)^{-1}) \\ &\leq R_{ni} + 2k^{1/(2k)} I(i(n)^{-1})^{1/(2k)}, \end{aligned}$$

wobei $R_{ni} := q^{-1}(\bar{H}_n(t_i^n)) |U_n(t_{i+1}^n-) - U_n(t_i^n)|$.

Daraus erhalten wir nun die Behauptung von 4.2.2, denn es gilt (man beachte (*) im Beweis zu 4.2.1):

$$\begin{aligned} \mathbb{E}(\{\max\{R_{ni}: i \in E\}\}^{4k}) &\leq \sum_{i \in E} \mathbb{E}(R_{ni}^{4k}) \\ &\leq \sum_{i \in E} C(k, m) 2i(n)^{-2k} q^{-4k}(\bar{H}_n(t_i^n)) \\ &\leq 2C(k, m) \left[\sum_{i \in E} [\bar{H}_n(t_i^n) - \bar{H}_n(t_{i-1}^n)]^k q^{-2k}(\bar{H}_n(t_i^n)) \right]^2 \\ &\leq 2C(k, m) k^2 I(\delta)^2. \end{aligned}$$

□

Analog zu 4.2.2 kann man zeigen:

4.2.3 Korollar. Es existiert ein C (nur abhängig von m) so, daß für alle $0 \leq a < b \leq 1$, $0 < \varepsilon$ und $n \in \mathbb{N}$ gilt:

$$\mathbb{P}(\sup\{|U_n(t) - U_n(s)| : a \leq \bar{H}_n(s) \leq \bar{H}_n(t) \leq b\} \geq \varepsilon)$$

$$\leq C \varepsilon^{-4} [(b-a)^2 + i(n)^{-2}].$$

Aus 4.2.3 folgt unmittelbar (man vgl. dieses Ergebnis mit 3.2.6):

4.2.4 Lemma. Es existiert eine nur von m abhängige Konstante C so, daß für alle $0 < \delta \leq 1$, $0 < \epsilon$ und $n \in \mathbb{N}$ mit $i(n) \geq \delta^{-1}$ gilt:

$$\mathbb{P}(\sup\{|U_n(t) - U_n(s)| : |\bar{H}_n(t) - \bar{H}_n(s)| \leq \delta\} \geq \epsilon) \leq C \epsilon^{-4} \delta.$$

4.3 L_p -Konvergenz des gewichteten U-Prozesses

Wie in 4.2 gehen wir wieder davon aus, daß ein $k \in \mathbb{N}$ existiert mit der Eigenschaft: q erfüllt (Q_k) .

Somit sind die Resultate 4.2.2 bzw. 4.2.4 anwendbar, die den Schlüssel darstellen für die Herleitung der L_p -Folgenkompaktheit von $(q^{-1}(\bar{H}_n)U_n)_{n \in \mathbb{N}}$. Natürlich benötigen wir dazu wieder eine zu (A) in 2.3 analoge Bedingung, nämlich:

(D) Es existiert eine Halbmetrik s auf $\mathbb{R}$ mit:

$$(i) \lim_{n \rightarrow \infty} \sup \sup\{|\bar{H}_n(t) - \bar{H}_n(r)| : s(r, t) \leq \delta; r, t \in \mathbb{R}\} \downarrow 0 \text{ für } \delta \downarrow 0.$$

(ii) $(\mathbb{R}, s)$ ist totalbeschränkt.

Man sieht sofort (vgl. hierzu 2.2.6), daß (D) insbesondere erfüllt ist (mit $s(r, t) := v((-\infty, r] \Delta (-\infty, t])$), falls ein finites Maß v auf $\mathbb{R}$ existiert mit:

$$(D') \sup\{\lim_{n \rightarrow \infty} \sup\{|\bar{H}_n(t) - \bar{H}_n(r)| : v((-\infty, r] \Delta (-\infty, t]) \leq \delta; r, t \in \mathbb{R}\} \downarrow 0 \text{ für } \delta \downarrow 0.$$

Dabei ist klar, daß (D') eine sehr viel allgemeinere Bedingung darstellt als (D) (i) (dafür ist in (D) die Halbmetrik i.a. nicht induziert durch ein finites Maß v).

Wir werden die folgenden L_b -Aussagen stets unter der Bedingung (D) herleiten und dabei im Modell von Kapitel I formulieren, wobei (in Anlehnung an die dortigen Bezeichnungen) $X=\mathbb{R}$ und s wie in (D). Dann lassen sich o.E. die $q^{-1}(\bar{H}_n)U_n$, $n \in \mathbb{N}$, auffassen als Zufallselemente in $\ell^\infty(\mathbb{R})$ (man beachte, daß dies mit Wahrscheinlichkeit 1 zutrifft). Mit $S=\ell^\infty(\mathbb{R})$ sowie $A=\{\pi_t(S): t \in \mathbb{R}\}$ sind dann trivialerweise auch die Meßbarkeitsvoraussetzungen aus Kapitel I gesichert. Falls nun (D) erfüllt ist, bezeichne

$$\omega_n(\delta) := \sup\{|q^{-1}(\bar{H}_n(t))U_n(t) - q^{-1}(\bar{H}_n(r))U_n(r)| : s(r,t) \leq \delta\}$$

für $n \in \mathbb{N}$ und $0 < \delta \leq 1$ den Stetigkeitsmodul des gewichteten U-Prozesses (bzgl. der Halbmetrik s aus (D)).

Aus den Ergebnissen des Abschnitts 4.2 erhalten wir nun unmittelbar:

4.3.1 Theorem. Sei (D) erfüllt und $i(n) \rightarrow \infty$ für $n \rightarrow \infty$. Dann ist $(q^{-1}(\bar{H}_n)U_n)_{n \in \mathbb{N}}$ relativ L_b -folgenkompakt.

Beweis. Aus (D) (i) ergibt sich sofort für alle $\epsilon > 0$:

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(\omega_n(\delta) \geq \epsilon) \leq \lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(\tilde{\omega}_n(\delta) \geq \epsilon), \text{ wobei}$$

$$\tilde{\omega}_n(\delta) := \sup\{|q^{-1}(\bar{H}_n(t))U_n(t) - q^{-1}(\bar{H}_n(r))U_n(r)| : |\bar{H}_n(t) - \bar{H}_n(r)| \leq \delta\}$$

für $0 < \delta \leq 1$ und $n \in \mathbb{N}$.

Wegen 4.2.2 und 4.2.4 folgt daraus (man beachte, daß aus Symmetriegründen eine zu 4.2.2 analoge Abschätzung auch für " $\bar{H}_n$ nahe bei 1" gilt):

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(\omega_n(\delta) \geq \epsilon) = 0 \text{ für alle } \epsilon > 0, \text{ d.h. (iii) (A)}$$

aus 1.3.4 ist erfüllt ((iii) (B) ist trivial). $\square$

Analog zum Beweis von 3.3.3 erhält man aus 4.3.1:

4.3.2 Theorem. Sei (D) erfüllt und $i(n) \rightarrow \infty$ für $n \rightarrow \infty$. Dann sind äquivalent:

- (i) $(q^{-1}(\bar{H}_n)U_n)_{n \in \mathbb{N}}$ ist L_b -konvergent gegen einen zentrierten Gaußprozeß G.
- (ii) Die Kovarianzfunktion von $(q^{-1}(\bar{H}_n)U_n)_{n \in \mathbb{N}}$ konvergiert (und zwar gegen die von G).

4.3.3 Bemerkung. Theorem 4.3.2 stellt eine Verallgemeinerung des Hauptergebnisses in SILVERMAN(1983) (dort Theorem A) in dreierlei Hinsicht dar: Zum ersten werden dort keine Dreiecksschemata betrachtet, zum zweiten wird nur die "i.i.d."-Situation abgehandelt (die hier natürlich enthalten ist), und zum dritten ist unsere Integralbedingung allgemeiner als die bei Silverman. Dazu betrachten wir zunächst:

- (i) Für $\epsilon > 0$ und $M \geq 1$ gilt: $(Q_{M+\epsilon})$ ist allgemeiner als (Q_M) , wobei wir die Definition von (Q_k) , $k \in \mathbb{N}$, auf alle $R \geq 1$ übertragen. Es gilt nämlich:

$$\begin{aligned} \int_0^{1/2} q^{-2(M+\epsilon)}(x) x^{M+\epsilon-1} dx &= \int_0^{1/2} q^{-2M}(x) x^{M-1} [q^{-2M}(x) x^M]^{\frac{\epsilon}{M}} dx \\ &\leq \int_0^{1/2} q^{-2M}(x) x^{M-1} \left[M \int_0^x q^{-2M}(u) u^{M-1} du \right]^{\frac{\epsilon}{M}} dx \\ &\leq M^{\frac{\epsilon}{M}} \left[\int_0^{1/2} q^{-2M}(x) x^{M-1} dx \right]^{1+\epsilon/M}. \end{aligned}$$

$(Q_{M+\epsilon})$ ist sogar echt allgemeiner als (Q_M) . Dazu betrachte

man die Funktion

$f(x) := x^{1/2} [\log \frac{1}{x}]^{1/(2M)}$ für $0 \leq x \leq \exp(-1)$ (dort ist diese Funktion strikt wachsend) und gemäß der Forderungen aus 4.1 geeignet auf $[0,1]$ fortgesetzt.

$$\begin{aligned} \text{Dann gilt: } \int_0^{\exp(-1)} f^{-2M}(x) x^{M-1} dx &= \int_0^{\exp(-1)} x^{-1} [\log \frac{1}{x}]^{-1} dx \\ &= \int_1^{\infty} u^{-1} du = \infty. \end{aligned}$$

Andererseits folgt:

$$\begin{aligned} \int_0^{\exp(-1)} f^{-2M-2\varepsilon}(x) x^{M+\varepsilon-1} dx &= \int_0^{\exp(-1)} x^{-1} [\log \frac{1}{x}]^{-1-\varepsilon/M} dx \\ &= \int_1^{\infty} u^{-1-\varepsilon/M} du < \infty. \end{aligned}$$

(ii) Die Bedingung (Q_1) ist bereits echt allgemeiner als die Silvermansche Bedingung an die Gewichtsfunktion q .

Dort wird nämlich gefordert:

$$(Q') \quad \int_0^{1/2} [\log \frac{1}{x}]^{1/2} v(dx) < \infty, \text{ wobei } v(x) := x^{1/2} q^{-1}(x)$$

wachsend ist in x auf $[0, \frac{1}{2}]$ mit $v(x) \rightarrow 0$ für $x \rightarrow 0$.

Mittels partieller Integration ergibt sich daraus:

$$\begin{aligned} \int_0^{1/2} [\log \frac{1}{x}]^{1/2} v(dx) &\stackrel{(*)}{=} v(\frac{1}{2}) [\log 2]^{1/2} + \int_0^{1/2} v(x) [2x (\log \frac{1}{x})^{1/2}]^{-1} dx \\ &\geq \int_0^{1/2} q^{-2}(x) [2v(x) (\log \frac{1}{x})^{1/2}]^{-1} dx \\ &\stackrel{(*)}{\geq} \int_0^{\varepsilon} q^{-2}(x) dx \quad \text{für genügend kleines } \varepsilon > 0. \end{aligned}$$

Dabei haben wir Gebrauch gemacht von der nachstehenden Eigenschaft:

$$(*) \quad v(x) \left[\log \frac{1}{x} \right]^{1/2} \longrightarrow 0 \text{ für } x \rightarrow 0, \text{ denn für alle } 0 < x \leq \frac{1}{2}$$

$$\text{gilt: } v(x) \left[\log \frac{1}{x} \right]^{1/2} \leq \int_0^x \left[\log \frac{1}{u} \right]^{1/2} v(du).$$

Damit ist also gezeigt, daß (Q_1) allgemeiner ist als (Q') .

Um nun noch "echt" allgemeiner nachzuweisen, betrachten wir die folgende Funktion:

$$f(x) := \left[x \log \frac{1}{x} \right]^{1/2} \left[\log \log \frac{1}{x} \right]^{1/2 + \delta} \text{ für } 0 < \delta \leq \frac{1}{2},$$

$0 < \varepsilon \leq e$ ($\varepsilon \leq \exp(-e)$) so gewählt, daß f strikt monoton

und f gemäß der Forderungen aus 4.1 geeignet auf

$[0,1]$ fortgesetzt.

Dann gilt (mit $\alpha = \log \log e^{-1}$; $\beta = \alpha^{-1/2 - \delta}$; $v(x) = x^{1/2} f^{-1}(x)$):

$$\int_0^\varepsilon f^{-2}(x) dx = \int_\alpha^\infty u^{-1-2\delta} du \leq (2\delta)^{-1} < \infty, \text{ wohingehend}$$

$$\int_0^\varepsilon \left[\log \frac{1}{x} \right]^{1/2} v(dx) = \beta + \int_0^\varepsilon v(x) \left[2x \left(\log \frac{1}{x} \right)^{1/2} \right]^{-1} dx$$

$$= \beta + \int_\alpha^\infty \frac{1}{2} u^{-1/2 - \delta} du = \infty \text{ (wegen } \delta \leq \frac{1}{2}).$$

(iii) Insgesamt folgt aus (i) und (ii), daß die von uns betrachtete Funktionenklasse sehr viel reichhaltiger ist als die in SILVERMAN(1983) zugelassene. Zudem fällt die Monotoniebedingung an v weg. Abschließend sei an dieser Stelle noch erwähnt, daß der Beweis zu 4.3.2 sehr viel einfacher ist als der bei Silverman (dort wird KMT verwendet).

4.3.4 Bemerkung. Mit Blick auf 3.3 stellt sich die Frage, ob in 4.3.1 $\bar{H}_n$ nicht durch die Varianz des U-Prozesses ersetzt werden kann. Dies setzt jedoch i.a. ein Abschneiden des Prozesses voraus. Wir wollen nun fordern, daß ein $K > 0$ existiert sowie eine Folge $(R_n)_{n \in \mathbb{N}}$ von monoton wachsenden, rechtsseitig stetigen Funktionen von $\mathbb{R}$ nach $[0,1]$ mit der Eigenschaft:

$$\mathbb{E}([U_n(t) - U_n(s)]^2) \leq K[R_n(t) - R_n(s) + i(n)^{-1}] \text{ für } s < t.$$

Man beachte, daß wegen 3.2.9 diese Bedingung insbesondere erfüllt ist mit $R_n := \bar{H}_n$. Wir untersuchen nun den Prozeß

$$W_n(t) := \begin{cases} q^{-1}(R_n(t)) U_n(t) & \text{falls } i(n)^{-1} \leq R_n(t) \leq 1 - i(n)^{-1} \\ 0 & \text{sonst,} \end{cases}$$

wobei q eine Gewichtsfunktion ist wie in 4.1, für die wir als zusätzliche Integrierbarkeitsbedingung verlangen:

$$\int_0^{1/2} q^{-2k}(x) x^{k-2} dx < \infty \text{ für ein } k \in \mathbb{N} \text{ mit } k \geq 2.$$

Wie in 4.1 zerlegen wir nun $\mathbb{R}$ für jedes $n \in \mathbb{N}$ wie folgt:

$$s_{i+1}^n := \inf\{s > s_1^n : R_n(s) - R_n(s_1^n) \geq i(n)^{-1}\}, \text{ wobei } s_0^n := -\infty \text{ und}$$

$\inf\{\emptyset\} := \infty$. Aus der vorausgesetzten Monotonie und Beschränktheit von R_n ergibt sich sofort, daß ein r_n existiert mit $s_{r_n}^n = \infty$. Sei weiter $(t_i^n)_{i \leq m_n}$ die Zerlegung aus 4.1 (bzgl. $\bar{H}_n$).

$$\text{Mit } \mathfrak{I}(s) := \int_0^s q^{-2k}(x) x^{k-2} dx \text{ für } 0 \leq s \leq \frac{1}{2} \text{ erhalten wir dann}$$

analog zu 4.2.1:

$$(1) \quad \mathbb{P}(\sup\{|W_n(s_1^n)| : R_n(s_1^n) \leq \delta\} \geq \varepsilon) \leq C \varepsilon^{-4k} \mathfrak{I}(\delta)^2 \text{ für alle } 0 < \delta \leq 1/2, 0 < \varepsilon \text{ und } n \in \mathbb{N}.$$

Weiter folgt für alle $n \in \mathbb{N}$ und $i \leq r_n$ mit $s_i^n < \infty$ (wobei C eine variable, von n, i, ϵ unabhängige Konstante ist):

$$\begin{aligned}
 (2) \quad & \mathbb{P}(\sup\{|U_n(t) - U_n(s_i^n)| : s_i^n \leq t < s_{i+1}^n\} \geq 3\epsilon) \\
 & \leq \mathbb{P}(\sup\{|U_n(t_1^n) - U_n(t_r^n)| : \mathbb{E}(|U_n(t_1^n) - U_n(t_r^n)|^2) \leq \frac{C}{i(n)}\} \geq \epsilon) \\
 & \quad + 2\mathbb{P}(\sup\{|U_n(t) - U_n(s)| : |\bar{H}_n(t) - \bar{H}_n(s)| \leq \frac{C}{i(n)}\} \geq \epsilon) \\
 & \stackrel{3.2.8}{\leq} C \epsilon^{-4k} i(n)^{2-2k} \quad \text{für alle } \epsilon > 0.
 \end{aligned}$$

Analog zum Beweis von 4.2.2 folgt dann aus (1) und (2):

$$(3) \quad \mathbb{P}(\sup\{|W_n(t)| : R_n(t) \leq \delta\} \geq \epsilon) \leq C \epsilon^{-4k} \tilde{I}(\delta)^2 \quad \text{für alle } 0 < \delta \leq 1/2, n \in \mathbb{N}.$$

Wir betrachten folgende Bedingung:

(E) Es existiert eine Halbmetrik s auf $\mathbb{R}$ mit:

$$(i) \quad \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup\{|R_n(t) - R_n(u)| : s(u, t) \leq \delta\} = 0.$$

(ii) $(\mathbb{R}, s)$ ist totalbeschränkt.

Dann folgt analog zu 4.3.1:

Theorem. Falls (E) erfüllt ist und $i(n) \rightarrow \infty$ für $n \rightarrow \infty$, ist

$(W_n)_{n \in \mathbb{N}}$ relativ L_b -folgenkompakt (im üblichen Modell bzgl. s aus (E)) und es sind äquivalent:

(i) $(W_n)_{n \in \mathbb{N}}$ ist L_b -konvergent gegen einen zentrierten Gaußprozeß G .

(ii) Die Kovarianzfunktion von $W_n, n \in \mathbb{N}$, konvergiert.

Das folgende Beispiel zeigt, daß dieses Ergebnis manchmal zu einer stärkeren Gewichtung führt als 4.3.1 (dort ist jedoch kein "Abschneiden" der Prozesse nötig).

4.3.5 Beispiel. Seien $\xi_{ni}, i \leq i(n), n \in \mathbb{N}$ über $[0,1]$ gleichverteilte zufällige Variable. Sei weiter $m=2$ sowie $h: [0,1] \xrightarrow{2} \mathbb{R}, (x,y) \mapsto \max(x,y)$ und $q(t) = [t(1-t)]^{1/3}$ für $0 \leq t \leq 1$. Dann gilt: es existiert $K > 0$ mit

$$\mathbb{E}([U_n(t) - U_n(s)]^2) \leq K(t^3 - s^3 + i(n)^{-1}) \quad \text{sowie}$$

$$\bar{H}_n(t) = t^2 \quad \text{für alle } 0 \leq s \leq t \leq 1 \text{ und } n \in \mathbb{N}.$$

Mit $R_n(t) := t^3$ erhalten wir nun gemäß Theorem in 4.3.4 die L_b -Konvergenz von $(W_n)_{n \in \mathbb{N}}$, wobei $W_n, n \in \mathbb{N}$, wie in 4.3.4 definiert ist bzgl. q und R_n . "Nahe bei 0" liefert dies also eine Gewichtung $q^{-1}(R_n(t)) \geq t^{-1}$, ein Ergebnis, daß wir aus 4.3.2 nicht gewinnen können (dazu wäre als Gewichtsfunktion $q'(t) \leq t^{1/2}$ "nahe 0" nötig; eine solche Funktion erfüllt jedoch die Bedingung $(Q_k), k \in \mathbb{N}$, nicht).

4.4 Raten im Glivenko-Cantelli für den U-Prozeß auf der reellen Achse im i.i.d.-Fall

In diesem Abschnitt wollen wir kein Dreiecksschema mehr zugrundelegen, sondern eine Folge $(\xi_i)_{i \in \mathbb{N}}$ von unabhängigen, identisch nach F verteilten zufälligen Variablen betrachten. Ansonsten übernehmen wir alle Voraussetzungen aus 4.1 sowie die dortigen Definitionen vermöge der Festsetzung: $\xi_{ni} := \xi_i$ für $i=1, \dots, n$ und $n \geq m$. Unser Ziel wird es nun sein, die $\mathbb{P}$ -fast sichere Konvergenz von Ausdrücken der Form $a_n D_n^q$ zu untersuchen, wobei für $n \geq m$

$$D_n^q := \sup \{ q^{-1}(H(t)) | H_n(t) - H(t) | : 0 < H(t) < 1 \}, \quad H := \bar{H}_m$$

und $(a_n)_{n \geq m}$ eine monoton wachsende Folge positiver Zahlen.

Ganz entscheidend für unser Vorgehen wird sein, daß wegen

3.1.4 für $n \geq m$ gilt:

$$\begin{aligned} (n) \quad \mathbb{E}(D_n^q) &\leq \frac{1}{n!} \sum_{j=1}^n \mathbb{E}(\sup(\|\frac{n}{m}\|^{-1/2} q^{-1}(H(t)) | \beta_n^j(t) | : 0 < H(t) < 1)) \\ &\leq \mathbb{E}(\bar{D}_n^q) . \end{aligned}$$

Dabei ist für alle $n \in \mathbb{N}$ und $j \leq n!$ β_n^j ein empirischer Prozeß zu $\|\frac{n}{m}\|$ unabhängigen, nach H verteilten Variablen ($\|\cdot\|$ bezeichne die Gaußklammer), α_n der uniforme empirische Prozeß sowie $\bar{D}_n^q := \sup\{n^{-1/2} q^{-1}(t) | \alpha_n(t) | : 0 < t < 1\}$. Weiter nehmen wir wieder o.E. (da dies mit Wahrscheinlichkeit 1 der Fall ist) an, daß D_n^q bzw. $\bar{D}_n^q$ zufällige Variable sind (d.h. endlich).

Wir erhalten zunächst:

4.4.1 Lemma. Falls für ein $0 \leq c \leq 1$ gilt: $\int_0^1 q^{-1-c}(x) dx < \infty$,

existiert eine Nullfolge $(\gamma_n)_{n \in \mathbb{N}}$ sowie ein $R > 0$ mit:

$$\mathbb{E}(n^{-1/2} S_n^q) \leq R n^{-c/(1+c)} \gamma_n^{1-c} \quad \text{für alle } n \in \mathbb{N},$$

wobei $S_n^q := \sup\{q^{-1}(t) | \alpha_n(t) | : 0 < t \leq \frac{1}{2}\}$.

Beweis. Mit SCHNEEMEIER(1983), Kapitel II, Lemma 5

folgt sofort (da α_n stationäre Zuwächse besitzt):

$$\begin{aligned} \mathbb{P}(R_n^q \geq \varepsilon) &\leq \varepsilon^{-2} \left(\frac{1}{2} - \frac{1}{n}\right)^{-2} \int_0^{1/2-1/n} q^{-2}(x+\frac{1}{n}) dx \\ &\leq 36 \varepsilon^{-2} \int_{1/n}^{1/2} q^{-2}(x) dx \quad \text{für alle } n \geq 3. \end{aligned}$$

Dabei ist $R_n^q := \sup\{q^{-1}(t) |\alpha_n(t) - \alpha_n(\frac{1}{n})| : \frac{1}{n} \leq t \leq \frac{1}{2}\}$.

Durch Integration über ε erhalten wir für alle $n \geq 3$:

$$\begin{aligned} \mathbb{E}(R_n^q) &\leq \int_0^\infty \mathbb{P}(R_n^q \geq \varepsilon) d\varepsilon \\ &\leq \int_0^\infty \min\{1, 36 \varepsilon^{-2} \int_{1/n}^{1/2} q^{-2}(x) dx\} d\varepsilon \\ &= \left[\int_{1/n}^{1/2} q^{-2}(x) dx \right]^{1/2} \underbrace{\int_0^\infty \min\{1, 36 u^{-2}\} du}_{=: C_1 < \infty} \\ &=: C_1 < \infty \end{aligned}$$

Mit $I(s) := \int_0^s q^{-1-c}(x) dx$ für $0 \leq s \leq \frac{1}{2}$ folgt daraus für $n \geq 3$:

$$\begin{aligned} \mathbb{E}(n^{-1/2} R_n^q) &\leq C_1 n^{-1/2} \left[q^{-1}(\frac{1}{n}) \int_{1/n}^{1/2} q^{-1-c}(x) dx \right]^{1/2} \\ &\leq C_1 n^{-c/(1+c)} \left[n^{-1} q^{-1-c}(\frac{1}{n}) \right]^{(1-c)/(2+2c)} I(\frac{1}{2})^{1/2} \\ &\leq C_1 n^{-c/(1+c)} I(\frac{1}{n})^{(1-c)/(2+2c)} I(\frac{1}{2})^{1/2}. \end{aligned}$$

Wegen $S_n^q \leq R_n^q + T_n^q$ mit $T_n^q := \sup\{q^{-1}(t) |\alpha_n(t)| : 0 \leq t \leq \frac{1}{n}\}$ für

alle $n \in \mathbb{N}$ müssen wir also lediglich eine geeignete Abschätzung für T_n^q finden, um den Beweis zu beenden. Wie im Beweis zu 4.2.2 (dort die Abschätzung für S_1) folgt sofort:

$$\begin{aligned} \mathbb{E}(n^{-1/2} T_n^q) &\leq 2 \int_0^{1/n} q^{-1}(x) dx \\ &\stackrel{\text{Hölder}}{\leq} 2 \left[\int_0^{1/n} dx \right]^{c/(1+c)} \left[\int_0^{1/n} q^{-1-c}(x) dx \right]^{1/(1+c)} \\ &= 2 n^{-c/(1+c)} I(\frac{1}{n})^{1/(1+c)} \quad \text{für } n \geq 2. \quad \square \end{aligned}$$

Mit der Momentenbedingung 4.4.1 sind wir nun in der Lage, Glivenko-Cantelli-Aussagen für den U-Prozeß herzuleiten. Dabei machen wir auch sehr wesentlich Gebrauch von der Tatsache, daß $(D_n^q)_{n \in \mathbb{N}}$ ein inverses Submartingal ist (vgl. Lemma 5 in GÄNSSLER(1983) in der Situation $m=1$; der Beweis für beliebiges m verläuft analog). Durch Anwendung der Chowschen Ungleichung (vgl. GÄNSSLER(1977), 6.6.1) erhalten wir:

4.4.2 Theorem. Für ein $0 \leq c \leq 1$ möge gelten: $\int_0^1 q^{-1-c}(x) dx < \infty$.

Für jede monoton wachsende Folge $(a_n)_{n \geq m} \subset (0, \infty)$ mit

$$\sum_{n=m+1}^{\infty} (a_n - a_{n-1}) n^{-c/(1+c)} < \infty \text{ und}$$

(i) $c < 1$ und $a_n = o(n^{c/(1+c)})$ oder

(ii) $c=1$ und $a_n = o(n^{1/2})$ gilt:

$$a_n D_n^q \rightarrow 0 \text{ } \mathbb{P}\text{-f.s.}$$

Beweis. Mit Hilfe der Chowschen Ungleichung folgt für jedes $0 < \varepsilon, m \leq j < 1$:

$$\varepsilon \mathbb{P}(\max(a_n D_n^q : n=j, \dots, 1) \geq \varepsilon) \leq \sum_{n=j+1}^1 (a_n - a_{n-1}) \mathbb{E}(D_n^q) + a_j \mathbb{E}(D_j^q)$$

$$\leq \sum_{n=j+1}^1 (a_n - a_{n-1}) \mathbb{E}(\bar{D}_{\lfloor \frac{n}{m} \rfloor}^q) + a_j \mathbb{E}(\bar{D}_{\lfloor \frac{j}{m} \rfloor}^q)$$

$$\stackrel{4.4.1}{\leq} 2R \left[\sum_{n=j+1}^1 (a_n - a_{n-1}) n^{-c/(1+c)} \gamma_n^{1-c} + a_j j^{-c/(1+c)} \gamma_j^{1-c} \right].$$

Dabei sind R und $(\gamma_n)_{n \in \mathbb{N}}$ wie in 4.4.1 gewählt. Für alle

$\varepsilon > 0$ gilt also: $\varepsilon \mathbb{P}(\max(a_n D_n^q : n \geq j) \geq \varepsilon) \rightarrow 0$ für $j \rightarrow \infty$. $\square$

4.4.3 Korollar. Sei $f: (0,1] \rightarrow (0,\infty)$ eine monoton fallende Funktion mit $\int_0^1 f(x)dx < \infty$. Falls dann ein $0 \leq c \leq 1$ existiert mit $\int_0^1 q^{-1-c}(x)dx < \infty$, folgt:

$$f(n^{-c/(1+c)}) D_n^q \rightarrow 0 \text{ } \mathbb{P}\text{-f.s.}$$

Beweis. Dies ist eine unmittelbare Folgerung aus 4.4.2,

wenn wir setzen: $a_n := f(n^{-c/(1+c)})$ für $n \geq m$. $\square$

4.4.4 Korollar. Es gelte: $\int_0^1 q^{-1-c}(x)dx < \infty$ für ein $0 \leq c \leq 1$.

Dann folgt:

$$q^{-1-c}(\frac{1}{2}n^{-c/(1+c)}) D_n^q \rightarrow 0 \text{ } \mathbb{P}\text{-f.s.}$$

Beweis. Ergibt sich sofort aus 4.4.3 mit $f(x) := q^{-1-c}(\frac{x}{2})$.

Aus 4.4.3 erhalten wir insbesondere:

4.4.5 Korollar.

(i) Falls $\int_0^1 q^{-1}(x)dx < \infty$, folgt: $D_n^q \rightarrow 0 \text{ } \mathbb{P}\text{-f.s.}$

(ii) Falls $\int_0^1 q^{-2}(x)dx < \infty$, folgt für alle $\epsilon > 0$:

$$n^{1/2} [(\log n)(\log \log n)^{1+\epsilon}]^{-1} D_n^q \rightarrow 0 \text{ } \mathbb{P}\text{-f.s.}$$

Beweis. (i) ist eine Anwendung von 4.4.3 für $f=1, c=0$.

(ii) ergibt sich ebenfalls aus 4.4.3 mit:

$$f(x) := [x \log \frac{1}{x} (\log \log \frac{1}{x})^{1+\epsilon}]^{-1} \text{ (ist monoton bei 0), } c=1. \square$$

4.4.6 Bemerkung.

- (i) Die Ergebnisse dieses Abschnitts lassen sich auch auf Dreiecksschemata übertragen, sofern man nur an IP-stochastischen Aussagen interessiert ist.
- (ii) 4.4.5(i) ist identisch mit Theorem 2.2 in HELMERS-JANSSEN-SERFLING(1985).
- (iii) Die Raten in 4.4.2 können verglichen werden mit Theorem (9) in NOLAN-POLLARD(1985) (dort werden allerdings nach Funktionenklassen indizierte U-Prozesse untersucht). Unsere Gewichtsfunktionen sind häufig besser, da für manche Kerne h die dortige Größe $v(\cdot)$ immer von 0 weg beschränkt ist (also gar keine Gewichtung stattfindet), wohingegen dies für H nicht der Fall ist.

4.4.7 Beispiel. Wir wollen zeigen, daß 4.4.5(i) i.a.

nicht verschärft werden kann. Dazu betrachten wir eine Folge $(\xi_i)_{i \in \mathbb{N}}$ von unabhängigen, über $[0,1]$ gleichverteilten Variablen und wählen als Kernfunktion

$$h(x_1, \dots, x_m) := \max(x_1, \dots, x_m) \text{ für } (x_1, \dots, x_m) \in [0,1]^m.$$

Dabei ist m fest gewählt und wir benutzen die Definitionen aus 4.4. Für beliebiges $n \geq m$ gilt dann:

$$(+)\quad H_n(t) = R_n \prod_{i=0}^{m-1} (F_n(t) - \frac{i}{n}) \text{ für } 0 \leq t \leq 1, \text{ wobei } F_n \text{ die}$$

empirische Verteilungsfunktion bezeichne zu $\xi_1, \dots, \xi_n$

$$\text{und } R_n := n^m \frac{(n-m)!}{n!} \geq 1.$$

Wir setzen nun voraus: $\int_0^{1/2} q^{-1}(x) dx = \infty$, wobei q ansonsten

die Bedingungen aus 4.1 erfülle.

Zu beliebigem $M \geq q^{-1}(\frac{1}{2})$ und $n \in \mathbb{N}$ sei $0 < a_n \leq \frac{1}{2}$ definiert durch $q^{-1}(a_n^m) = M n^m$ (im Fall $m=1$ ist dieses Vorgehen identisch zum Beweis von Theorem 1(ii) in GÄNSSLER(1983), S. 20).

Dann ist $(a_n)_{n \in \mathbb{N}}$ monoton fallend und es gilt:

$$\begin{aligned} \sum_{n=1}^{\infty} a_n^m n^{m-1} &\geq 2^{-m} \sum_{n=1}^{\infty} a_n^m [(n+1)^m - n^m] \\ &= 2^{-m} M^{-1} \sum_{n=1}^{\infty} a_n^m [q^{-1}(a_{n+1}^m) - q^{-1}(a_n^m)] \\ &\geq 2^{-m} M^{-1} \int_0^{a_1^m} q^{-1}(x) dx = \infty. \end{aligned}$$

Mit MORI(1976), Lemma 3, folgt daraus sofort:

(*) $\mathbb{P}(\xi_{n:m} \leq a_n \text{ für unendlich viele } n \in \mathbb{N}) = 1$, wobei für $n \in \mathbb{N}$ $\xi_{n:m}$ die m-te Ordnungsgröße ist von $\xi_1, \dots, \xi_n$.

Falls nun jedoch für ein $n \in \mathbb{N}$ gilt: $\xi_{n:m} \leq a_n$, erhalten wir sofort wegen (+):

$$q^{-1}(H(a_n)) H_n(a_n) = q^{-1}(a_n^m) H_n(a_n) \geq R_n M m! \geq M.$$

Mit (*) folgt $\mathbb{P}(\limsup_{n \rightarrow \infty} \sup\{q^{-1}(H(t)) H_n(t) : 0 < t \leq \frac{1}{2}\} \geq M) = 1$,

also mit $M \uparrow \infty$:

$$(++) \mathbb{P}(\limsup_{n \rightarrow \infty} \sup\{q^{-1}(H(t)) H_n(t) : 0 < t \leq \frac{1}{2}\} = \infty) = 1.$$

Wir unterscheiden nun zwei Fälle:

Fall 1: $\sup\{q^{-1}(H(t)) H(t) : 0 < t \leq \frac{1}{2}\} < \infty$. Dann erhalten wir mit (++):

$$\text{sofort: } \mathbb{P}(\limsup_{n \rightarrow \infty} D_n^q = \infty) = 1.$$

Fall 2: $\sup\{q^{-1}(H(t))H(t): 0 < t \leq \frac{1}{2}\} = \infty$. Dann gilt, wobei

$\xi_{n:1}$ die erste Ordnungsgröße von $\xi_1, \dots, \xi_n$ bezeichne für $n \in \mathbb{N}$:

$$\mathbb{P}(D_n^q = \infty) \geq \mathbb{P}(\sup\{q^{-1}(H(t))H(t): 0 < t < \xi_{n:1}\} = \infty) = 1,$$

$$\text{also auch } \mathbb{P}(\limsup_{n \rightarrow \infty} D_n^q = \infty) = 1.$$

Insgesamt erhalten wir also in beiden Fällen, daß

$$\limsup_{n \rightarrow \infty} D_n^q = \infty \text{ } \mathbb{P}\text{-f.s.}$$

LITERATURVERZEICHNIS

- ALEXANDER, K. (1984a): Probability inequalities for empirical processes and a law of the iterated logarithm. Ann. Prob. 12, 1041-1067.
- ALEXANDER, K. (1984b): Central limit theorems for stochastic processes under random entropy conditions. University of Washington. Preprint.
- ALEXANDER, K. (1985): The central limit theorem for weighted empirical processes indexed by sets. University of Washington. Preprint.
- BILLINGSLEY, P. (1968): Convergence of probability measures. New York: Wiley.
- DUDLEY, R.M. (1978): Central limit theorems for empirical measures. Ann. Prob. 6, 899-929.
- GÄNSSLER, P. und STUTE, W. (1977): Wahrscheinlichkeitstheorie. Berlin: Springer (Hochschultext).
- GÄNSSLER, P. und STUTE, W. (1979): Empirical processes: a survey of results for independent and identically distributed random variables. Ann. Prob. 7, 193-243.
- GÄNSSLER, P. (1983): Empirical processes. IMS Lecture Notes-Monograph Series, Vol. 3.
- GÄNSSLER, P., HÄUSLER, E. und SCHNEEMEIER, W. (1984): Selected topics on empirical processes. In: Proceedings of the Third Prague Symposium on Asymptotic Statistics. Ed. by P. Mandl and M. Hušková.- Amsterdam-New York-Oxford: Elsevier Science Publishers B.V., p. 57-91.
- GÄNSSLER, P. (1985): Bootstrapping empirical measures indexed by Vapnik-Chervonenkis classes of sets. Preprint.
- GINE, E. und ZINN, J. (1984): Some limit theorems for empirical processes. Ann. Prob. 12, 929-989.
- HELMERS, R., JANSEN, P. und SERFLING, R. (1985): Glivenko-Cantelli properties of some generalized empirical distribution functions and strong convergence of generalized L-Statistics. Department of Mathematic Sciences. The John Hopkins University, Baltimore, Maryland. Preprint.

- HOEFFDING, W. (1963): Probability inequalities for sums of bounded random variables. J. Amer. Stat. Assoc. 58, 13-30.
- KIEFER, J. (1961): On large deviations of the empiric D.F. of vector chance variables and a law of iterated logarithm. Pacific J. Math. 11, 649-660.
- MORI, T. (1976): The strong law of large numbers when extreme terms are excluded from sums. Z. Wahrsch. Verw. Gebiete 36, 189-194.
- NOLAN, D. und POLLARD, D. (1985): U-Processes: Rates of convergence. Statistics Department Yale University. Preprint.
- POLLARD, D. (1981): Limit theorems for empirical processes. Z. Wahrsch. verw. Gebiete 50, 317-331.
- SCHNEEMEIER, W. (1983): Invarianzprinzipien für gewichtete empirische Prozesse und Anwendungen. Diplomarbeit am Mathematischen Institut der Ludwig-Maximilians-Universität München.
- SERFLING, R. (1984): Generalized L-, M- and R-Statistics. Ann. Stat. 12, 76-85.
- SHORACK, G.R. (1979): The weighted empirical process of row independent random variables with arbitrary distribution functions. Statistica Neerlandica 33, No. 4, 169-189.
- SILVERMAN, B.W. (1983): Convergence of a class of empirical distribution functions of dependent random variables. Ann. Prob. 11, 745-761.

LEBENS LAUF

Geburtsdatum: 19.12.1956

Geburtsort: Wertingen

Schulbildung: September 1963 bis August 1967
Volksschule Wertingen;
September 1967 bis Juni 1976
Johann-Michael-Sailer-Gymnasium
Dillingen/Donau;
Reifeprüfung: 25. Juni 1976

Bundeswehr: Juli 1976 bis September 1977

Studium: Mathematik an der Ludwig-Maximilians-Universität München vom
WS 1977/78 bis SS 1983

Abschlußprüfungen: Diplomprüfung in Mathematik
am 13. Mai 1983

Berufliche Tätigkeiten: Seit 1. Juli 1983 Wissenschaftliche
Hilfskraft am Mathematischen Institut der Ludwig-Maximilians-Universität München

